THE AMERICAN
JOURNAL OF PSYCHOLOGY

EDITED BY

KARL M. DALLENBACH
UNIVERSITY OF TEXAS

AND

M. E. BITTERMAN   E. B. NEWMAN
INSTITUTE FOR ADVANCED STUDY   HARVARD UNIVERSITY

WITH THE COÖPERATION OF

CONTENTS

The American Journal of Psychology

(Contents—continued)

L. L. Thurstone

THE DIRECT ESTIMATION OF SENSORY MAGNITUDES—LOUDNESS

By S. S. STEVENS, Harvard University

The purpose of this study was to try to develop and refine a method for the direct quantitative assessment of subjective magnitudes. It has been amply demonstrated in recent years that people can say when one stimulus appears half as great as a standard stimulus, or when one stimulus appears twice as great as a standard.[1] Ratios other than 2:1 have also been tested with fairly consistent results.[2] In these experiments the ratio is usually predetermined by the experimenter (E), and the observer's (O's) task is to find or produce the stimuli that arouse two sensory magnitudes answering to the prescribed ratio.

To this procedure there is, of course, an inverse: we can preset the stimuli and ask O to specify the apparent ratio between them. The development and exploration of this inverse process, which we might call the *method of magnitude-estimation*, has been the object of the present series of studies.

Historically, the production of stimuli to fit a prescribed subjective ratio appears to have been tried first by Merkel.[3] This was in 1888 in his *Methode der doppelten Reize*. He worked with brightness, loudness, and finger pressure, and tried to find

---

[1] D. W. Robinson, The relation between the sone and the phon scales of loudness, *Acustica*, 3, 1953, 344-358.

[2] P. H. Geiger and F. A. Firestone, The estimation of fractional loudness, *J. acoust. Soc. Amer.*, 5, 1933, 25-30; R. M. Hanes, The construction of subjective brightness scales from fractionation data: a validation, *J. exp. Psychol.*, 39, 1949, 719-728.

[3] J. Merkel, Die Abhängigkeit zwischen Reiz und Empfindung, *Phil. Stud.*, 4, 1888, 541-594; 5, 1889, 245-291; 5, 1889, 499-557.

the stimulus that appeared to double the sensation. A prototype of the inverse experiment was carried out by Richardson and Ross, who presented two different sound-intensities and asked O to estimate the apparent ratio between their loudnesses, calling the first or standard loudness 1.00.[4] In these early experiments the engineering problems involved in producing and measuring the stimulus were usually so difficult that few useful results were obtained. Ham and Parkinson later used a slightly different version of this method: O was presented a standard followed by a comparison tone of reduced intensity and was told "to judge what per cent of the original loudness was left."[5] Unfortunately, it appears that the different comparison levels were always presented in descending order, which may have distorted the estimates made at the lower intensities.

In my own experiments on loudness, brightness, and lifted weights, I have tried several versions of the method of magnitude-estimation in an effort to appraise the assets and liabilities of the procedure. It all started from a friendly argument with a colleague, who said, "You seem to maintain that each loudness has a number and that if someone sounded a tone I should be able to tell him the number." I replied, "That's an interesting idea. Let's try it." We agreed that, as in any problem of measurement, we would need first to decide on a modulus—a size for our unit—so I sounded a loud tone and we agreed to call its loudness 100. I then presented a series of other intensities in random order, and, with a readiness that surprised both of us, he assigned numbers to them in a thoroughly consistent manner.

That was my first use of the method. Only after working with this procedure for a couple of months did I discover—or rediscover—that it is basically similar to the method used by Richardson and Ross, which I had described in 1938.[6] How easily one forgets!

Anyhow, the evidence accumulated over the past two years suggests that, if properly used, the method of magnitude-estimation can provide a simple, direct means of determining a scale of subjective magnitude. The method has wide potential utility, but like all psychophysical methods it has its pitfalls and its sources of potential bias. In any given situation, most of these distorting factors can probably be discovered and either avoided or balanced out in the experimental design.

By a scale of subjective magnitude we mean a quantitative scale by which we can predict what people will say when they try to give a quantitative description of their impressions. Since we are concerned with how people judge stimuli, we must rely on one form or another of verbal report. Consequently, the construction of a representative scale of sensation involves a process not unlike public-opinion polling, and as every pollster knows there are good ways and bad ways of asking questions. The problem in subjective measurement is to arrange the conditions and the task in such a way that O can assess his impressions and communicate them to E in a quantitative language with as few biasing cues, suggestions, and constraints as possible. The fact that a person can be influenced in what he reports does not mean that he has no

[4] L. F. Richardson and J. S. Ross, Loudness and telephone current, *J. gen. Psychol.*, 3, 1930, 288-306.

[5] L. B. Ham and J. S. Parkinson, Loudness and intensity relations, *J. acoust. Soc. Amer.*, 3, 1932, 511-534.

[6] S. S. Stevens and Hallowell Davis, *Hearing: Its Psychology and Physiology*, 1938, 113.

impressions, or that they are not quantifiable, but only that the task is a difficult one, *O* behaves like a sensitive galvanometer. He is sensitive not only to the stimuli he is trying to gauge but also to a host of adventitious influences that in varying degrees can warp his reactions. Nevertheless, under optimal conditions the typical individual is able to make direct ratio estimations of relative subjective magnitude over wide ranges—as much as a billion to one (90 db.) in physical intensity, which corresponds approximately to a thousand to one in subjective ratio. These estimates provide important corroboration of the subjective scales generated by other procedures such as fractionation. In particular they demonstrate clearly that loudness is a power function of the stimulating intensity.

An Illustrative Experiment

Let us first consider an illustrative experiment. The stimulus-tone was 1000~ whose intensity was controlled by *E*. *O* sat before a pair of push-type switches, one of which served to produce the standard loudness and the other the variable loudness. To suppress switching transients, the tone was passed through a narrow band-pass filter after which it went to a pair of earphones (PDR-8's mounted in MX-41/AR cushions).

The task can best be described by means of the instructions:

Instructions. The left key presents the standard tone and the right key presents the variable. We are going to call the loudness of the standard 10 and your task is to estimate the loudness of the variable. In other words, the question is: if the standard is called 10 what would you call the variable? Use whatever numbers seem to you appropriate—fractions, decimals, or whole numbers. For example, if the variable sounds 7 times as loud as the standard, say 70. If it sounds one fifth as loud, say 2; if a twentieth as loud, say 0.5, etc.

Try not to worry about being consistent; try to give the appropriate number to each tone regardless of what you may have called some previous stimulus.

Press the 'standard' key for 1 or 2 sec. and listen carefully. Then press the 'variable' for 1 or 2 sec. and make your judgment. You may repeat this process if you care to before deciding on your estimate.

The standard level was set at a sound-pressure level (*SPL*) of 80 db. (*re* 0.0002 dyne/cm^2). There were 9 variable levels covering a range of 70 db. and spaced as shown by the data points in Fig. 1 (circles). The series of variables was presented to each of 18 *O*s twice in irregular order in a single session. The order of presentation was different for each *O* and each of the nine variables served for at least one *O* as the first to be estimated. Six of the 18 *O*s had not previously observed in this type of experiment; the other 12 had been used in experiments involving somewhat different procedures. This was the first experiment in which the *O*s had ever judged stimuli presented at levels both above and below the standard.

As shown by the upper curve in Fig. 1, the median judgments show remarkable consistency. The numerical estimates fall close to a straight line in a log-log plot, which indicates that subjective loudness is a power function of physical intensity (acoustic power). The straight line in Fig. 1

is actually a segment of the loudness scale determined from the pooled data of the previous experiments in this field.[7] The empirical formula relating loudness L to intensity I for a 1000 ~ tone turns out to be

$$L = kI^{0.3}.$$

In this equation I refers to energy flux density and is assumed to be pro-
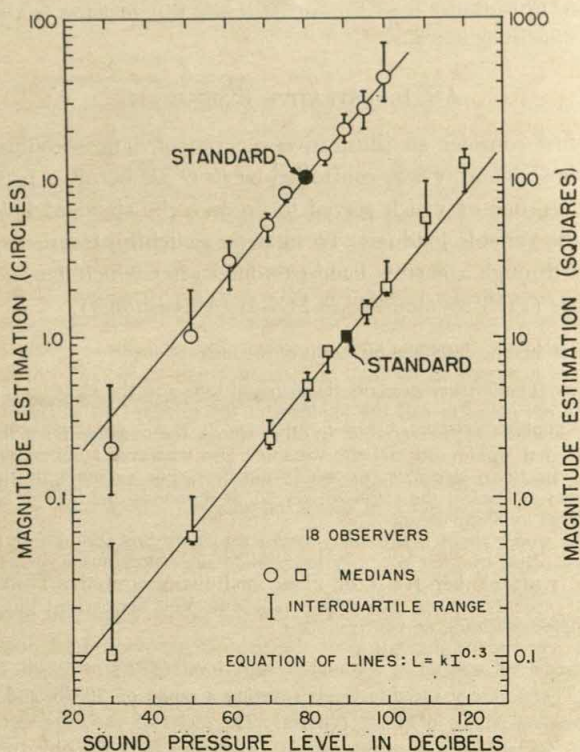
FIG. 1. MAGNITUDE-ESTIMATIONS: UPPER CURVE: STANDARD SET AT 80 DB. AND CALLED 10. LOWER CURVE: STANDARD SET AT 90 DB. AND CALLED 10.
Points are the medians for two estimations by each of 18 Os. The vertical lines mark the interquartile ranges. The straight lines through the standards are segments of the 'S' scale of loudness for tones of 1000~.

portional to the square of the sound pressure. This is the equation of the line in Fig. 1. Since this is not quite the same as the loudness scale

[7] Stevens, The measurement of loudness, *J. acoust. Soc. Amer.*, 27, 1955, 815-829.

commonly used in the past, I will refer to this function as the 'S' loudness scale.[8]

The variablity of the magnitude-estimations is indicated in Fig. 1 by the vertical lines marking the interquartile ranges. As we might expect, the variability increases as the loudness ratio between the standard and the variable grows larger. An approximate empirical formula relating the interquartile range of the estimates to the loudness ratio being judged can be written as

$$\log Q_3/Q_1 = 0.4 \quad |\log L_V/L_S|,$$

where Q_3 and Q_1 are the two quartile points, L_V and L_S are the loudnesses of the variable and the standard, and the vertical bars indicate absolute value. In terms of stimulus-intensities this formula becomes:

$$\log Q_3/Q_1 = 0.12 \quad |\log I_V/I_S|.$$

The distributions of the estimates are nearly always skewed toward the higher numbers, as is indicated by the fact that the means are nearly always larger than the medians. Because of this skewness it is more appropriate to use medians than means. There are always a few Os whose estimates deviate far from those of the majority, but the behavior of the typical (median) individual is usually quite consistent.

I have reported the foregoing experiment first on the list because I believe it is a fair example of how well the method of magnitude-estimation can be made to work. It demonstrates that normal, intelligent Os can make consistent quantitative appraisals of their subjective experiences. On the other hand, not all of our experiments have turned out so well. There are numerous pitfalls in this method—as in any method that attempts to assess human judgments—and our main concern in what follows will be to try to tease out some of the factors that can influence the results.

First let me say that the success of the foregoing experiment was achieved only after much trial and error in the course of which we learned at least some of the things *not* to do. On the positive side, in experiments involving a fixed standard, some of the things that should be done appear to be these:

(1) Use a standard whose level does not impress the O as being either extremely soft or extremely loud, *i.e.* use a comfortable standard—one that O can 'take hold of.'

(2) Present variable stimuli that are both above and below the standard.

(3) Call the standard by a number, like 10, that is easily multiplied and divided.

(4) Assign a number to the standard *only,* and leave the O completely

---

free to decide what he will call the variables. For example, do not tell O that the faintest variable is to be called '1,' or that the loudest is to be called some other number. If E assigns numbers to more than one stimulus, he introduces constraints of the sort that force O to make judgments on an interval rather than on a ratio scale.

(5) Use only one level of the standard in any one session, but use various standards, for it is risky to decide the form of a magnitude function on the basis of data obtained with only one standard.

(6) Randomize the order of presentation. With inexperienced Os it is well, however, to start with loudness ratios that are not too extreme and are, therefore, easier to judge.

(7) Make the experimental sessions short—about 10 min.

(8) Let O present the stimuli to himself. He can then work at his own pace, and he is more apt to be attending properly when the stimulus comes on.

(9) Since some estimates may depart widely from those of the 'average' O, it is advisable to use a group of Os that is large enough to produce a stable median.

These bits of advice are probably not all of equal importance. Some of them may be necessary; some of the others may be merely desirable. It would be a tedious task indeed to try to assess all these factors by putting them through all possible experimental permutations.

As we shall see later, good results can apparently be obtained even when no standard is designated, in which case some of the foregoing rules obviously do not apply.

Effect of Range and Level of the Stimulus

In the illustrative experiment discussed above, the variable stimuli were spread out over a range of 70 db., corresponding to a subjective loudness ratio of about 200 to 1. The consistency of the results led to a more ambitious experiment involving a stimulus-range of 90 db. The question was: could O, in one sitting, consistently estimate the loudness of tones that ranged from very faint to uncomfortably loud? For most of the Os the lowest variable stimulus, at 30 db. *SPL*, seemed extremely faint, and the loudest stimulus, at 120 db. *SPL*, always produced startle and wincing when it first came on. The standard was again called 10, but its level was set at 90-db. *SPL*. The 18 Os comprised 10 from the previous experiment and 8 new ones. Each O made two judgments of each variable in random order.

The results are shown by the lower curve in Fig. 1. According to these

magnitude estimates, the most intense tone was about 1000 times louder than the faintest tone. Over the middle range of intensity the median values of the subjective estimates fall close to the 'S' loudness scale which, as explained above, was derived from previous experiments. At the extremes of the range there is, however, a systematic departure from the loudness function. The loudness of the intense tones is overestimated; the loudness of the faint tones is underestimated. This underestimation of the faint tones is also evident in the results of the previous experiment (upper curve of Fig. 1).

Of course, in speaking of over- and under-estimation we are presuming to know what the 'true' curve is like, which may or may not be our good fortune. It might be argued, for example, that a curved line through the points of Fig. 1 would be a better loudness function than the straight line drawn. This might indeed be plausible were it not for the evidence of numerous other experiments, some of which we shall now consider. This evidence will show that the steepness of the function caused by the over- and under-estimations in Fig. 1 can be reversed by the simple expedient of changing the level of the standard stimulus. In other words, the too-steep aspect of the ends of the observed curve in Fig. 1 can be made to appear too flat by setting the standard at one or the other extreme of the range.

The data in Fig. 2 illustrate this point. The experimental procedure was the same as before. When the standard was at 120 db. (and called 100) the median estimates determined a curve whose slope is less than the expected loudness-function—instead of greater as in Fig. 1. The same tendency is observed in the data obtained with the standard at 30 db. (and called 1). Thus, we have altered the slope of the observed function by setting the level of the standard at one or the other extreme of the range.

In another pair of experiments (same procedure) the standards were set at 100 db. (called 50) and at 50 db. (called 1). The results, shown in Fig. 3, are precisely what we should expect. Both above and below the 100-db. standard the loudness is overestimated. With the standard at 50 db., which is a moderately faint level, the situation is reversed: both above and below the standard the loudness is underestimated.

These six experiments taken together suggest that the 'true' loudness-scale is probably a power function—a straight line in a log-log plot. At least, we see that deviations in both directions from this straight-line function can be produced by altering the level of the standard in appropriate ways.

One is tempted, of course, to try to explain these changes in the magnitude estimations by *ad hoc* hypotheses, but the difficulty is that, if any one hypothesis will do the job, it is probable that many other hypotheses would do as well. Nevertheless, I venture to state that the simplest explanation is as follows: In making these estimations the O does two things: he estimates the variable relative to the standard—as he is supposed to do—but he also weights his judgment to a slight extent by a factor related to the absolute level of the variable tone. In other words, he seems to be trying to say, for example, that such-and-such variable is not only five times louder than the standard but it is also a very loud tone indeed. So he estimates the variable as perhaps six times louder than the standard. Then, in the reverse situation, when

the standard is made very loud, O's reaction seems to be that a particular variable may be five times fainter than the standard, but since it is still quite loud he estimates it as perhaps only four times fainter. Of course, O is probably not conscious of this process, nor is he conscious of the reverse process, which occurs when the tones are absolutely very faint.

If this tendency to flavor the relative judgment by a slight admixture of an abso-
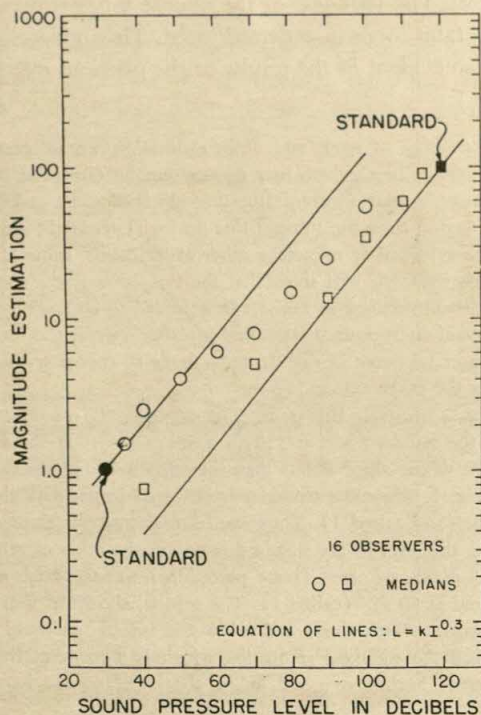
FIG. 2. MAGNITUDE ESTIMATIONS: STANDARD AT 30 DB. CALLED 1; AT 120 DB.
CALLED 100
With these extreme standards the slopes determined by the data are flatter than the 'S' scale of loudness.

lute judgment is in fact operating in these experiments, the bias produced can in principle be balanced out by combining the evidence from experiments like those illustrated in Fig. 2 with those in Fig. 1. The principle involved here is not unlike that which leads the chemist to interchange the scale-pans when he performs a precision weighing.

EFFECT OF PRESENTING THE STANDARD ONCE ONLY

When these studies were started, I was interested in discovering whether O could make consistent judgments of loudness without comparing each variable directly to

a standard. It turns out that he can—and, as we shall see later, he can even get along without a standard. Since some interesting facts emerge from experiments in which the standard is presented once only, I shall describe a typical study.

The standard (called 100) was presented only at the beginning of the experiment and then, without hearing the standard again, the O judged each of several variables presented in random order. Then, after a short rest, the standard was pre-
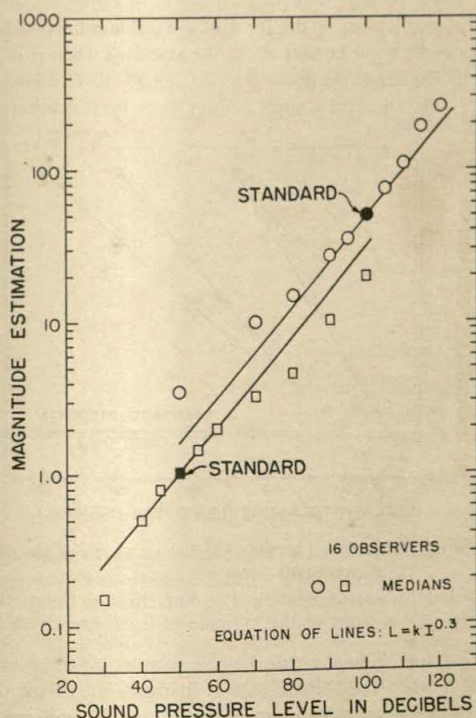
FIG. 3. MAGNITUDE ESTIMATIONS: STANDARD AT 50 DB. CALLED 1; AT 100 DB.
CALLED 50

Relative to the loud standard, the variable intensities are overestimated. Relative to the soft standard, they are underestimated.

sented again and the whole procedure was repeated. There was no systematic difference between the first and second runs. Three standards were used, 70 db. (16 Os), 100 db. (17 Os) and 120 db. (14 Os).

The medians of the direct estimates of the loudness of the variable tones are plotted in Fig. 4. It is quite apparent that even against a remembered standard the magnitude-estimates are linear in a log-log plot. It is also clear that under these conditions the level of the standard has a marked influence on the slope of the function. The straight line drawn through the 100-db. standard has the 'proper' slope—the slope of the 'S' scale of loudness. When the standard is at 70 db. the

slope is steeper; at 120 db. it is flatter. From the evidence of Fig. 1, as well as from that of several other experiments, we can predict that judgments based on the same sets of variable stimuli would have produced different slopes if the standards had been at the bottoms rather than at the tops of the ranges. Thus the too-flat slope in Fig. 4 between 70 and 120 db. would become a too-steep slope if the standard were placed at 70 db.

An interesting effect occurs when O is obliged to remember the standard for a period of time. His recollection of the level of the standard seems to shift systematically, as demonstrated by what he says about the standard when it is presented to him again at the end of the series. In about 9 cases out of 10 O reports that at the end of the series the 70-db. standard sounds fainter than he remembered it. Conversely,
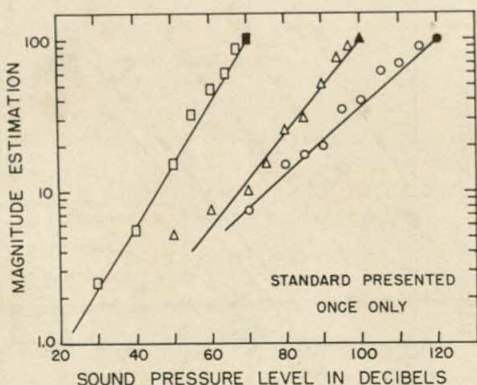
FIG. 4. MAGNITUDE ESTIMATIONS (MEDIANS) RELATIVE TO A STANDARD PRESENTED ONLY AT BEGINNING OF EACH SERIES

In each case the standard was called 100. The line through the 100-db. standard has the slope of the 'S' scale of loudness.

the 120-db. standard sounds louder than he remembered it—often much louder. This despite the fact that when the 120-db. tone is first presented most Os wince. Nevertheless, they seem to remember this uncomfortable stimulus as being less intense than it really was. As we might expect, there seems to be no systematic shift in the memory for tones in the middle range. We must conclude, therefore, that faint tones move up in memory, loud tones move down, and medium tones stay put.

This drift in itself would seem at first sight to account for the fact that some of the slopes in Fig. 4 are not the same as that of the 'S' scale of loudness. The drift in the memory of the standard undoubtedly contributes to the changes in slope, but there is probably more to it than that. A drift of the standard could hardly account for the slope changes in Fig. 1, for there we would need to assume that the drift occurs in two directions at once. We need at least the additional hypothesis that O weights his judgment by a factor related to the absolute level.

ESTIMATION VERSUS ADJUSTMENT

It may be of interest at this point to compare O's magnitude-estimations with the settings he makes when he adjusts a tone to a fractional loudness.

Fig. 5 shows this comparison. The data are from part of an experiment run by E. C. Poulton and the author. Four standards (called 100) were used: 60 db. (8 Os), 80 db. (16 Os), 100 db. (16 Os), and 120 db. (8 Os). The order in which the Os made the adjustments and the estimates was counterbalanced. In both cases O presented the stimuli to himself by pressing one or the other of a pair of keys as often as he wished. He made
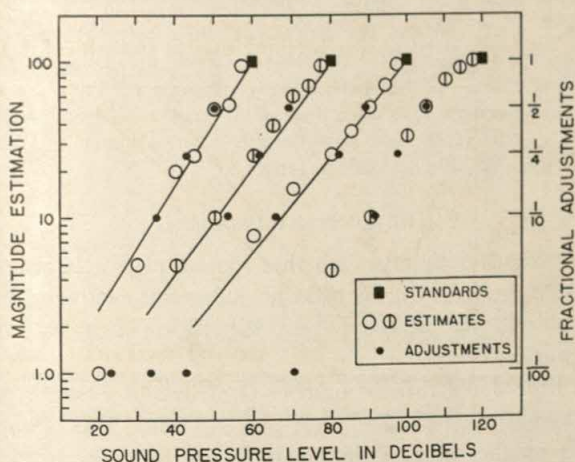
FIG. 5. MAGNITUDE ESTIMATIONS AND ADJUSTMENTS TO FRACTIONAL LOUDNESSES
The median magnitude estimates (left-hand ordinate) may be compared with the median values to which the Os adjusted the intensities to produce various fractional loudnesses (right-hand ordinate). The standards were called 100.

his adjustments with the aid of a 'sone potentiometer' whose characteristic was such that the subjective loudness it produced was roughly proportional to the position of the knob.[9]

First we note in Fig. 5 that the slope of the lines is, as usual, a function of the level of the standard. The line for the 100-db. standard has the slope of the 'S' scale of loudness. Again, we see evidence of a marked 'time error' when the standard is at 120 db. Thus, when the variable was 3 db. less than the standard, only one of the 8 Os failed to call the variable 100—meaning that it sounded as loud as the standard. When the variable was 6 db. less, 2 Os called it 100. This time-error, if that is what it should be called, seems to account for the marked curvature in the results for the 120-db. standard.

[9] S. S. Stevens and E. C. Poulton, The estimation of loudness by unpractised observers, *J. exp. Psychol.*, 51, 1956, 71-78.

The adjustments made to the fractional loudnesses at 0.5, 0.25, 0.1, and 0.01 agree well with the direct estimations except for the smallest fraction (0.01). For this fraction the Os do not reduce the intensity of the tone as much as their direct estimations would predict they should. We suspect that this may be due in part to a kind of conservative tendency which most Os exhibit when requested to adjust to extreme ratios. Part of the discrepancy may also be due to certain inadequacies in the apparatus.

In this same experiment we also used standards of 20, 40, 60, and 80 db. and asked the Os to make direct estimates of louder tones, and to make adjustments to the multiples 2, 4, 10, and 100. The data showed the expected change in slope, *e.g.* going up from the 80-db. standard the slope is steeper than going up from 20 db. We are not presenting the data here, however, because unfortunately we made a procedural mistake—as we have often done! This mistake involved forcing the O to use the very low end of the sone potentiometer where a minute turn of the knob makes a very large change in the loudness ratio.

THE CONTEXT PROBLEM

The question naturally arises whether context plays a large rôle in the estimation of magnitude. Does O make his judgments relative to the population of variable stimuli presented? Does the spacing of the variable stimuli make any difference?[10]

There is little doubt that the number O assigns to a tone is sometimes influenced by what he has called some previous tones. This is why it seems wise to mix up the order of the variables. On the other hand, we have strong evidence that the effect of context does not account for the basic fact that loudness is a power function of intensity.

Perhaps the most convincing evidence in support of this notion is the fact that the *first* variable heard by O is judged in a manner consistent with a power-function loudness-scale. When only one variable has been heard, it is hard to see how the other variables can have any influence on the judgment.

To test this proposition, we performed an experiment on a group of 32 undergraduate students, each of whom made only one magnitude-estimation.[11] Four groups of 8 Os each judged a different variable relative to a common standard of 100 db. (called 100). The median estimate of each group is plotted in Fig. 6. It is evident that, with no context supplied by a population of variable tones, the data fall fairly close to a straight line. On the first judgments ever made by these Os, they gave quantitative estimates of their sensations which agree remarkably well with the proposition that loudness is a power function of intensity.

[10] The large literature on contextual effects is too large to be evaluated here. Cf., however, Harry Helson, Adaptation-level as a basis for a quantitative theory of frames of reference, *Psychol. Rev.*, 55, 1948, 297-313.

[11] Stevens and Poulton, *op. cit.*, 77-78.

Under certain conditions, however, context can be made to play an important rôle, as Garner has already shown.[12] The effects of context can, however, be made negligible, or practically so, if the experiment is designed properly.

The spacing of the variable stimuli seems also to play only a negligible part in determining the estimations made in these experiments. Note, for example, how different the spacings of the stimuli are in Figs. 1-5. Actually, I had thought at first that spacing might be quite important. I began by spacing the variable stimuli evenly by 5-db. steps below the standard (called 100). Under these conditions, most of the
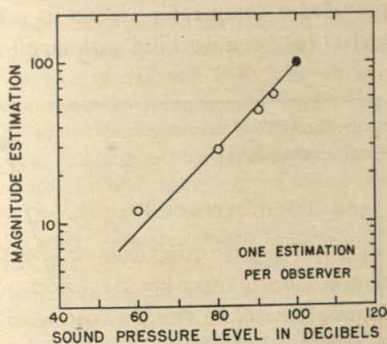
FIG. 6. MAGNITUDE ESTIMATES BASED ON SINGLE ESTIMATE BY EACH O. The standard at 100 db. was called 100. Each point is the median estimate for a group of 8 Os.

estimates are low numbers, below 50. It seemed wise, therefore, so to space the stimuli that O would give estimates that were more evenly spaced in terms of loudness. (Figs. 4 and 5 show this sort of spacing.) It appears, however, to make little, if any, difference what the spacing is—at least under these procedures, in which the O usually makes only two judgments of any one variable.

The results of a deliberate attempt to explore the matter of spacing are shown in Fig. 7. The experiment was conducted by Joseph C. Stevens. The crosses and the circles represent the mean results for two different experiments. In one the variable stimuli were spaced evenly by decibels (crosses); in the other, they were spaced to give more nearly uniform steps in loudness (circles). Each of 10 Os made two judgments of each variable in random order. Obviously, it seems to make little difference how the variable stimuli are spaced.

Incidentally, in Fig. 7 the ordinate has been made linear in order to show how these power functions look when a linear measure of loudness is plotted against a logarithmic measure like decibels. Fechner's law predicts that this curve should be a straight line. It is plainly far from straight. It should also be pointed out that, although the loudness function for white noise is similar to that for the 1000 ~ tones, there appear to be some slight differences between the two.

---

[12] W. R. Garner, Context effects and the validity of loudness scales, *J. exp. Psychol.*, 48, 1954, 218-224.

EFFECT OF CHANGING THE STANDARD

In most of our experiments O was given only one standard on any one day. At the end of a session, however, we have occasionally tried to change the level of the standard, but O usually finds this confusing. A tone he might have called 20 in the earlier session, he now must call something quite different.

It appears to be possible, however, to change the standard in the course of an experiment provided a comparable change is made in the number assigned to it. Thus, if O has been working with an 80-db. standard called
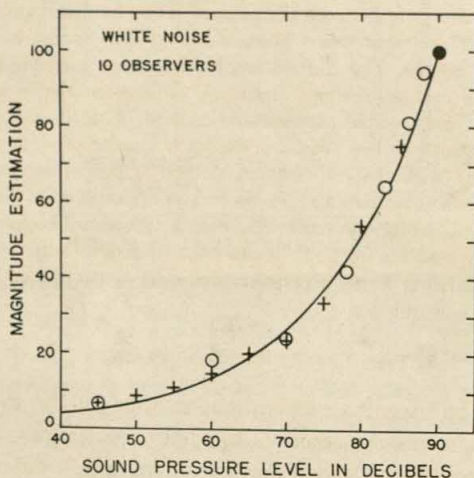
FIG. 7. THE EFFECT OF THE SPACING OF THE STIMULI
In one experiment the stimuli were spaced evenly by decibels (crosses). In the other they were spaced about evenly by loudness (circles). Each point represents the mean of 20 judgments—2 by each of 10 Os.

10 (as in Fig. 1) he can shift fairly easily to a 100-db. standard if it is called 50 (as in Fig. 3). He makes this shift easily because the new number assigned to the standard is about what he had been calling tones of this intensity when they were presented as variables.

We have also tried the extreme procedure of changing the standard on every presentation. Four standard levels were used, 60, 80, 100, and 120 db. The four standards were always called 100 and were paired with a variable of lower intensity. Then, in another experiment, the standards 30, 50, 70, and 90 db. were called 1, and they were also presented in random order—this time paired with a variable of greater intensity. The tones were

presented by a motor-driven timer so that O heard: standard (1 sec.), space (1 sec.), variable (1 sec.), space (2 sec.), standard (1 sec.), space (1 sec.), variable (1 sec.). He then made his judgment. Some of the 8 Os did quite well with this procedure, but others gave rather erratic results. All the Os reported that a constantly changing standard made the experiment very difficult. The Os find it easier to use one standard at a time because they can then settle on one modulus for their scale of loudness and make all their estimates in terms of a single 'subjective unit.'

This fact might well have a bearing on the utility of the so-called 'constant sum' method—an interesting procedure recently proposed by Metfessel.[13] Under this method each level presented is paired with every other level, and O is required to indicate the subjective ratio between the pairs of levels by distributing some number of points (e.g. 100) between them. Thus, if two stimuli appear to be in the ratio 4:1, O would say 80, 20. The method has recently been used for lifted-weights.[14] Compared with the method of direct magnitude estimation relative to a fixed standard, I would anticipate that the constant-sum method, with its ever changing standard and variable, would be less satisfactory from O's point of view.

Another difficulty with the constant-sum method is that it requires that O state the ratio between two subjective magnitudes and, at the same time, make the numerator and the denominator of the ratio add up to a constant. This seems like an unnecessary constraint to place on O. If he can estimate a ratio, why not let him report it simply and directly? It is my experience that every constraint we place on the O is a potential source of trouble.

THE EFFECT OF CONSTRAINTS

The method of magnitude-estimation, as used in the experiments reported thus far, seems to subject O to as few constraints as possible. In effect, all he is told is the size of the unit of the scale—the modulus—and beyond that he makes all the decisions. It is interesting to compare the typical results of this method with those of other experiments in which O is not so free.

Obviously the most extreme degree of constraint possible is to instruct O what number to assign to each tone. A version of this experiment was performed by M. S. Rogers in an attempt to see how well O could learn to identify different absolute loudnesses. The question was: how much information (in 'bits') can be transmitted by loudness? Rogers used 16 intensities spaced 6 db. apart. The three Os were told in advance that the problem was to identify the tones by number, 1 to 16. The tones were first sounded in ascending order and the appropriate number was assigned by

[13] M. F. Metfessel, A proposal for quantitative reporting of comparative judgments, J. Psychol., 24, 1947, 229-235; A. L. Comrey, A proposed method for absolute scaling, Psychometrika, 15, 1950, 317-325.
[14] J. P. Guilford and H. F. Dingman, A validation study of ratio-judgment methods, this JOURNAL, 67, 1954, 395-410.

E. Then the tones were presented in random order and *O* tried to identify the tone. After he made each guess he was told the correct designation.

Each *O* made at least 50 estimates for each intensity before the test-series was run. This test consisted of 10 judgments at each level. The average number assigned to the different intensities is shown by the circles in Fig. 8.

The fact that the average judgment falls close to the straight line demonstrates that *O*s can be constrained to give an arbitrary, predetermined designation to the loudnesses of a set of tones. With this degree of constraint operating we would
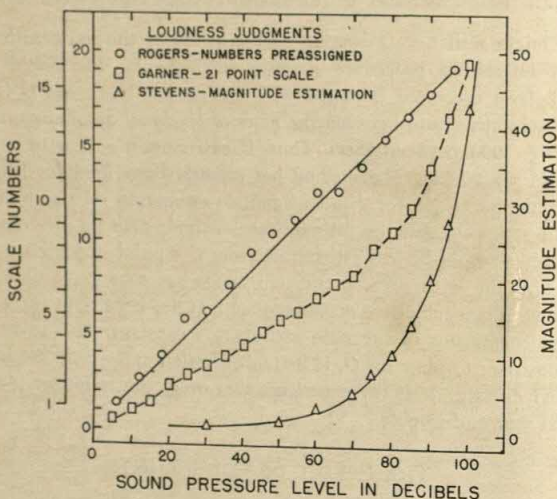
FIG. 8. COMPARISON OF EXPERIMENTS WITH THREE DEGREES OF CONSTRAINT
In Rogers' experiment (circles) the *O*s were taught to give a predetermined number (left-most scale) to each loudness heard. In Garner's experiment (squares) the *O*s were told to judge the loudness of 20 tones using *all* the numbers from 0 to 20. The magnitude estimations (triangles) are the same data as in Fig. 1 (upper curve). The curve through the triangles is the 'S' scale of loudness.

scarcely be tempted to call the resulting scale a valid measure of loudness, although it does obey Fechner's law. *O*s can give back Fechner's law in their magnitude-judgments provided they are taught to do so—but it takes a lot of teaching.

At the other extreme, under the method of magnitude-estimation, in which few constraints are operating, we obtain judgments that are related to intensity in the manner shown by the lower curve in Fig. 8. The data (triangles) are the same as those on the upper curve of Fig. 1. The smooth curve through the triangles is the 'S' scale of loudness. In Fig. 8 the judgment scale (ordinate) is linear instead of logarithmic.

Between these two extremes of minimal and maximal constraint, many degrees are, of course, possible. A good example of an intermediate degree of constraint can be found in an experiment that Garner conducted in the course of his studies of the problem of discriminability as it relates to the informational content of absolute

judgments.[15] Garner has kindly supplied me with some of his unpublished data, which I have plotted in Fig. 8. A median response was computed for each of 10 Os and these medians were averaged. Garner instructed his Os as follows.

You will hear a series of tones of different loudness presented one at a time every 6 sec. A light will flash 1 sec. before each tone is presented, so you will know when the tone is coming. Your task is to assign a number to each tone according to its apparent loudness to you. In this series the tones will vary from some that you can barely hear, or may not hear at all, to some very loud tones, and you should use numbers from 0 through 20. If you do not hear the tone at all, call it '0.' If you heard it, but it was as weak as any tone in this series, call it '1.' If it is as loud as any tone in this series, call it '20.' Use the other numbers to indicate the in-between loudnesses. For example, if a tone appears to be half-way between the softest and the loudest tones in the series, call it '10.' Remember to try to assign the numbers according to how loud the tones appear to *you*. We are interested in how loud tones *seem* to be to you, not in some kind of 'accuracy.' Use all of the numbers from 0 through 20 to indicate these loudnesses. First you will be given a few trials so you can establish the range of loudnesses you will be hearing.

There were 20 stimulus-intensities ranging from 5 to 100 db. and spaced 5 db. apart. Each stimulus was presented 100 times and the order of presentation of the stimuli was balanced.

The important point to note here is that O is asked to use the numbers to tell how loud the tones appear to be, *but* he is also told to use all 21 numbers (0–20) and to distribute them over the range of stimuli. If O's scale of loudness is anything like the lower curve in Fig. 8, then his task is obviously impossible. According to the scale of loudness, if the loudest tone is called 20, a tone 5 db. lower would be called about 14, a tone 10 db. lower would be called 10, and so forth. In the series of stimuli used by Garner, there is no item that ought properly to be called 17, for example. Under these circumstances, where O feels constrained to follow E's instructions to use all the numbers and at the same time to assess his experience of loudness, he does precisely what we should expect—he compromises. The result is a distribution of judgments shown by the squares in Fig. 8. This curve is about midway between the 'S' scale of loudness and the curve obtained in Rogers' experiment where O was under maximal constraint.

We might also note that, according to the scale of loudness, if the loudest tone used by Garner is called 20, his faintest tone should be called not 1 but something less than 0.1. Telling the O to call this tone 1 forces him to judge the remaining tones on what is, at best, an interval- rather than a ratio-scale.[16]

Another point should be made here. It applies to both Rogers' and Garner's experiments, as well as to some of the procedures we have tried out in magnitude-estimation. The point is this: if the O is told that the upper end of the scale is limited to some number or other, the *average* value assigned to the higher stimuli will be too low, because O can make errors in only one direction. Conversely, if a limit is placed on the lowest number to be assigned, the average of the numbers assigned at the low end of the scale will be raised. The net result of these effects will be to make the slope of the function less steep than it would otherwise be. This

[15] W. H. Garner and H. W. Hake, The amount of information in absolute judgments, *Psychol. Rev.,* 58, 1951, 446-459.

[16] Stevens, Mathematics, measurement and psychophysics, in S. S. Stevens (ed.) *Handbook of Experimental Psychology,* 1951, 27.

effect is usually small, but the fact that it exists at all is another reason for leaving O free to use any numbers that he thinks appropriate. It is also an argument in favor of placing the variable stimuli both above and below the standard—or at least letting O expect that the variable can be in either direction.

THE PROBLEM OF THE USE OF NUMBERS

As we have seen, the method of magnitude estimation requires that O use numbers to describe some aspect of his experience in quantitative terms. The method presupposes that O can make quantitative estimations in this manner. This, of course, is a lot to presuppose. It appears that most intelligent, educated people can make these quantitative estimates in a consistent manner, but it is idle to assume that all people can do so.

A discouragingly large segment of the population seems unable, for example, even to use numbers consistently on an election ballot that requires a rank ordering of the candidates. E. B. Newman and M. S. Rogers, in a study of the election ballots of the City of Cambridge, found that about 1500 ballots out of 37,000 were invalidated by an apparent inability of the voter to make use of numbers in one of the most elementary ways possible.

It is not surprising, therefore, that an occasional O in our experiments turns out to use numbers in curious ways. Even for the typical O, however, some ways of reporting magnitude-estimates are easier than others. In the experiments reported above, several different designations (moduli) have been given to the standard loudness, e.g. 1, 10, 50, and 100. All these moduli are probably easier to use than a modulus like 73, for example, but it turns out that these four are not equally well liked by the Os. The order of preference seemed to be 10, 100, 1, 50. When the standard is called 50 (see Fig. 3), O finds it more difficult to subdivide into such simple ratios as ⅓, ¼, etc.

Incidentally, experiments like these would probably be greatly facilitated if we did our counting on base 8, instead of on base 10. If our system of numeration repeated on base 8, instead of on base 10, successive halvings and doublings would be psychologically more easy,[17] and since many Os seem to use this process in making their magnitude-estimates, they would not, on base 8, be forced to deal with so many fractional values.

As we might expect, the Os show a preference for certain numbers. When the standard is called 100 most Os used numbers ending in 5 or 0. When the standard is called 10 and the variable is 50 db. below the standard most Os call it either 0.1 or 0.5. These preferences appear, however, to exert only a minor influence on the outcome.

We have also tried a radically different method of reporting magnitude estimations. This experiment was carried out on eight graduate students by E. C. Poulton. In one part of the experiment the standard (at 100 db.) was called 100 and O estimated the loudness of 5 fainter tones. In another part the standard (at 60 db.) was called 1 and O estimated the loudness of 5 louder tones. Then, in still another part of the experiment, the 100-db. standard was again called 1 but O was required

---

[17] Ibid., 6.

to estimate the loudness of the variable as a fraction with a variable denominator, *e.g.* 1/1.8, 1/8.5, etc. Half the *O*s started with the estimates based on the standard called 100; the other half started with the fractional estimates. The estimates involving multiples of 1 were always inserted between these two series.

The tones were presented by a motor-driven timer, as in the experiment discussed above in which the level of the standard was randomized. Each *O* judged each variable twice.

The median estimates obtained in these three experiments are shown in Fig. 9. For these 8 *O*s there certainly seems to be no significant difference between fractional estimates (triangles) and the estimates based on the modulus 100 (circles).
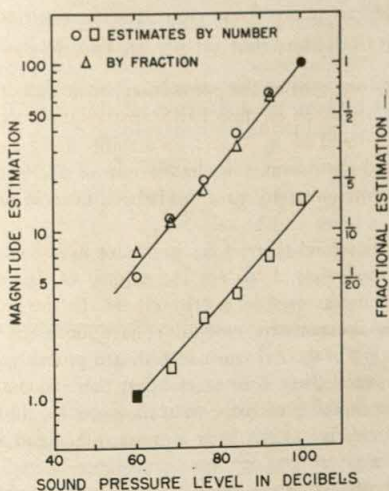
Fig. 9. Magnitude Estimations with Two Methods of Report
The 100-db. standard was called 100 (circles), or it was called 1. When it was called 1, the *O* reported his estimates in terms of a fraction with a variable denominator (triangles). The lower curve (squares) is for magnitude estimates, based on a standard of 60 db., called 1.

Both sets of estimates lie close to the line representing the 'S' loudness-scale.

Despite the agreement between the two sets of results, all 8 *O*s reported that the fractional estimates were much harder to make than the estimates based on the modulus 100.

The multiple estimates based on modulus 1 also seem to determine a straight line, but the slope is slightly less in this case.

ESTIMATIONS WITH NO DESIGNATED STANDARD

The evidence reviewed thus far suggests that what we call the standard and how *O* reports his judgments are not crucial to the general form of the loudness-function obtained. To carry this suggestion to its logical conclusion, another experiment was designed in which no standard was desig-

nated and no restriction was placed on $O's$ method of reporting, except that he was asked to use numbers proportional to loudness.

Eight tones spaced 10 db. apart from 40 to 110 db. were presented twice each in a different order to each of 26 Os. The first tone presented was varied from one O to another. The number of Os who started with each of the various levels were: 5 Os at 80 and at 90; 4 at 60 and at 70; 3 at 50 and at 100; and 1 at 40 and at 110. The instructions were as follows.

I am going to give you a series of tones of different intensities. Your task is to tell me how loud they sound by assigning numbers to them. To turn on the tone you simply press the key. You may press it as often as you like.

When you hear the first tone, give its loudness a number—any number you think appropriate. I will then tell you when to turn on the next tone, to which you will also give a number.

Try to make the ratios between the numbers you assign to the different tones correpond to the ratios between the loudnesses of the tones. In other words, try to make the numbers proportional to the loudness, as you hear it.

All but one of the Os had previously made estimations with the aid of a fixed standard, and these instructions came as a mild shock to some of them. Nevertheless, they set to work without much protest.

The designations given to the first tone presented were quite varied: the actual numbers the Os used were 1, 5, 8, 9, 10, 15, 20, 50, 90, and 100. The number 10 was the most popular; it was used by 8 Os.

The absolute size of the numbers used covers a wide range. Thus, of 2 Os who are always very consistent in their magnitude estimations, one called the loudest tone 18 and the softest 0.1, the other called the loudest 500 and the softest 5.

Since each O used a modulus of his own choosing, the problem of averaging the data presents some interesting angles. The procedure adopted was that of bringing the estimates made at a given intensity into coincidence by multiplying by an appropriate factor. Thus, the judgments made by each O were weighted by a factor that made the average of his two estimates at 80 db. turn out to equal 10. His other judgments were weighted in the same proportion. The median estimates were then computed and plotted in Fig. 10.

I also tried weighting the estimates so that the average estimates at 70 db. averaged out to 5. This procedure gave practically the same results as the first procedure, as shown by the median values recorded in Table I. The upper and lower quartile points are also entered in Table I.

The data presented in Fig. 10 bear a striking resemblance to the 'S' scale of loudness—the straight line in the figure. I will confess that I had scarcely expected such consistency, for the median results are, if anything, slightly more consistent than those obtained when the O could refer to a standard on each trial (Fig. 1). When there is no standard the first tone presented seems to become the standard, although it may remain so only temporarily. At the conclusion of the experiment, most of the Os reported that they felt uncertain whether they were able to remember the first tone presented. Actually, only five of the Os gave this first tone exactly the

same number when it came along the second time, but the largest ratio difference between the two assignments was a case in which 10 was assigned the first time and 30 the second time.

Since for most of the Os the experiment began with tones of medium intensity we find a slight tendency toward overestimation of the loudest tone and underestimation of the faintest tone, similar to that observed in Fig. 1. Perhaps this tendency would vanish if the starting levels were evenly distributed among the Os. On the other

TABLE I

Magnitude Estimations of Loudness with No Standard Stimulus

Twenty-six Os made two estimations of each stimulus. In the first case, the estimates of each O were multiplied by a factor to make the average estimate at 80 db. = 10. In the other case, the estimates were multiplied by a factor to make the average at 70 db. = 5.

Decibels SPL	80 db. = 10			70 db. = 5		
	Median	$\mathcal{Q}_3$	$\mathcal{Q}_1$	Median	$\mathcal{Q}_3$	$\mathcal{Q}_1$
110	83	110	57	77	143	50
100	40	57	29	42.5	58	22
90	20	25	15	19.5	25	13
80	10	12	8	10	12	7
70	5.4	6.3	4	5	6	4
60	2.8	3.6	2.2	2.5	3.5	2
50	1.4	1.8	0.8	1.2	2.0	0.8
40	0.6	0.9	0.4	0.5	0.8	0.3

hand, the Os seem to find the task more difficult when the first tone heard is near one or the other extreme of the range.

Apparently the Os do not dislike working with no standard, although most of them commented on the low subjective certainty they felt about their estimations. The typical comment was, "I must have been horribly inconsistent."

In this experiment, subjective certainty does not seem to have much to do with performance. On the other hand, some Os expressed a decided preference for the procedure in which no standard was used. One of them said:

I felt freer to use numbers over a wide range. I liked the idea that I could just relax and contemplate the tones. When there was a fixed standard I felt more constrained to try to multiply and divide loudnesses, which is hard to do; but with no standard I could just place the tone where it seemed to belong.

For some Os, then, even the designating of a standard by E constitutes an undesirable constraint. This I had not expected. Nor do I think it operates effectively as a constraint for all Os. Perhaps the most we can say is that the preferred procedure for one O is not the preferred procedure for all. It probably also depends on what experience O has had with judgments of this sort, and how familiar he is with loudness. Although I may be quite wrong, I imagine that starting from scratch with no standard would prove difficult for people who had never judged loudness before.

This final experiment seems to go about as far as an experiment can go toward getting the Os to make absolute judgments of loudness. Here we have done nothing more than present a random series of intensities, and
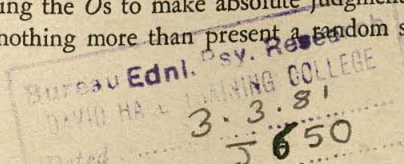
to each level O has assigned a number representing his perception of its loudness. The median responses fall close to the 'S' scale of loudness.

INTROSPECTIVE OBSERVATIONS

On many occasions the Os have been asked to try to say how they go about making their estimates. Some of them visualize a linear scale on
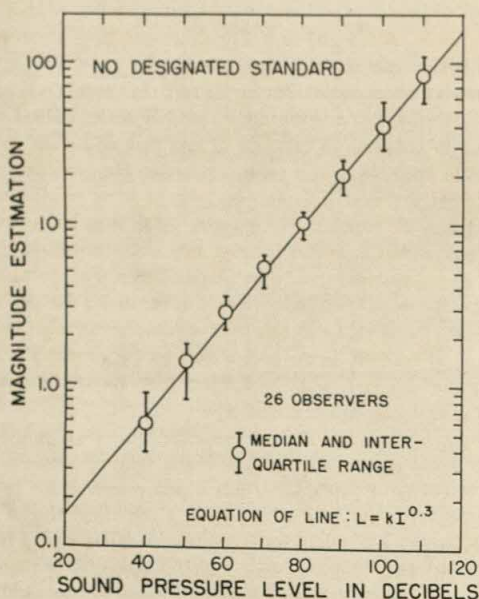
FIG. 10. MAGNITUDE ESTIMATIONS WITHOUT A FIXED STANDARD
Intensities were presented in irregular order and O used any numbers he chose in order to describe the loudnesses. Before the median was taken the numbers were multiplied by a factor to make each O's estimates for 80 db. average 10.

which they try to place the tones. Others say they have no visual imagery at all. Actually, some of those who say they cannot visualize are among the more atypical in their responses, but I cannot say this is a general rule.

As mentioned earlier, many Os find it difficult to use a very faint standard. They commonly state that they "just can't get hold of it" when its level is near threshold. Some Os come to the experiment with interesting preconceptions either about loudness or about numbers. For example, when using modulus 100, some Os do not give numbers below about 10 regardless of how faint the variable is made. When asked about this, some of them have said that they were afraid to "use up" the low numbers for fear something still fainter might come along. When E explained that no matter how low a number they might have used, there would still remain an

infinite number of numbers between this value and zero, they sometimes seemed surprised, but most of them saw the point when they thought about it.

Three Os out of about 30 recently tested stated that to them loudness-ratios were limited to about 10 or 20 to 1. They could not imagine a ratio as great as 100 to 1, or at least they were quite convinced that it would be silly to try to estimate it. When they were later given a tone of 120 db. followed by one of 30 db., they all admitted that the fainter one was less than a twentieth as loud—but two of them could not be induced to estimate the ratio. Interestingly enough, however, these same Os consistently made estimates of visual brightness over a subjective range of as much as 100 to 1. In the experiment with no standard, 2 of these 3 Os assigned numbers whose ratios exceeded 100 to 1. The other one gave numbers to the faintest and loudest tones in the ratio of only 14 to 1.

Another problem we encounter is due to the fact that some Os seem to make their estimates on an interval-scale, or even on an ordinal scale, instead of on the ratio-scale we are trying to get them to use. One O says that all he can do with the set of variables is to try to keep them in a consistent order. He gives one of the variables a number and then if the next is fainter he gives it some smaller number; if it is louder he gives it a larger number. He is pretty good at ordering the variables and seldom makes a reversal, but he is clearly using only an ordinal scale.

Two Os have said they have their own private scales and that 100 is what they would call the loudest tone possible. When I presented a 50 db. tone to one of these Os and called it 1, he said that this level was more like what he would call 20. He then proceeded, as he said, to assign numbers to the variables according to their position between the standard and the loudest tone he could imagine. In other words, he seemed to be using an interval scale.

Some Os have asked why I did not tell them the value they should call the standard and also the value they should assign to the faintest tone, or to the loudest tone, as the case might be. I explain that this would put us in the position of trying to determine an interval-scale, not a ratio-scale. I tell them that if I *knew* in advance what the faintest tone should be called, the problem would be solved before the experiment began. This is not an easy point for some Os to comprehend.

Despite these many difficulties encountered by some Os in making quantitative estimates of sensory magnitudes, the typical (median) O seems able to describe his sensations in a consistent quantitative language. At least he seems able to do so whenever E gives him a fair and unbiased opportunity to do so.

THE 'QUANTITY OBJECTION' AND THE 'STIMULUS-ERROR'

When I see with what assurance and consistency some Os (I might even say most Os) assess subjective magnitudes, I am puzzled that so much bitter ink should have been spilled over the so-called "quantity objection" to psychophysics.[18] Many of the older worthies tilted at this issue, and many of them denied the measurability of subjective magnitudes.

As one example, Fullerton and Cattell have this to say.

We can indeed say when one weight seems approximately double another, but this is doubtless because we have often lifted first one volume, and then two, and the

---

[18] E. G. Boring, The stimulus error, this JOURNAL, 32, 1921, 449-471.

like. But we cannot say when one sound seems twice as loud, or one day twice as hot as another.[19]

Ebbinghaus joins the objectors, but not to say that estimates cannot be made.[20] His objection is simply that when estimates are made they tell us only about the stimulus. Since he makes an interesting genetic argument here, I will quote it:

In general one designates the brightness of a flame or a surface as 10 or 12 times another brightness, and could just as easily, it appears, designate a loud tone as the double or treble of a soft tone. But what occurs here is no longer an immediate sensation or an immediate judgment of sensations, but depends upon the introduction of experiences. We can readily experience, and we do every day experience, the fact that the arousal of a brightness or a loudness depends upon a diversity of just those physical things or processes that in limited number call forth the impression of darker or softer. In order to have an impression of greater brightness for a surface, one can increase the number of gas flames illuminating it; in order to strengthen a tone, one multiplies the instruments carrying it. Such experiences with respect to the causes of sensations we have always in immediate view, and we believe that we have the numerical characteristics that always attach to the one occurring without anything further in the other.

This argument interests me particularly, because one of my early recollections is of an old-fashioned chandelier in our sitting room. The chandelier had two lamp bulbs, each with its own pull chain. When one chain was pulled the room lighted up to a fairly bright level. When the second chain was pulled, the brightness increased, but it certainly did not seem to double. This puzzled me. But Ebbinghaus says that we judge lights and sounds on the basis of experience of this sort. This also puzzles me.

The fact of the matter is that adding two equal lights or two equal sounds (in random phase) can do no more than increase the level by 3 db. With both light and sound a 3-db. increase in the stimulus is judged by the typical O to produce approximately a 25% increase in the subjective magnitude. It is far from the 100% increase that Ebbinghaus' genetic argument would lead us to expect.

We are forced to conclude, I think, that even though there are some situations in which we make the 'stimulus error' in judging light and sound our estimates are so far from reflecting the magnitude of the stimulus that the only reasonable hypothesis is that these estimates are valid measures of sensation itself.

THE PROBLEM OF VALIDITY

Even under the best conditions the quantitative estimation of subjective magnitudes is not an easy task. Sensations do not come with numbers written on them, and when we try to assess the ratio between a pair of them we find ourselves up against a difficult task of appraisal. It is no wonder then that subtle constraints and biases can influence the result. This is another way of saying that the outcome is a function of method—as it always is in science. What we want, of course, is an unbiased method, one that on the average lets O make an estimate that is neither too high

---

[19] G. S. Fullerton and J. McK. Cattell, *On the Perception of Small Differences,* 1892, 20.
[20] Boring, *op. cit.,* 454.

nor too low. Since we do not know in advance what his estimate should be, we can apply no independent criterion of validity.

Under these circumstances we are forced to pick a criterion on the basis of common sense. We have to make a value judgment—just as we have to make a value judgment when we decide that the use of a balance is the best means of setting up a fundamental scale of weight.

In the measurement of subjective magnitudes, it seems only reasonable to pick a procedure that 'makes sense' in terms of the problem before us: If the problem is to discover how the typical person assesses the loudness of sounds, we have every reason to accept the direct estimate he makes in a free situation—in a situation that leaves him free to select his quantifiers without constraint or prejudice. It is easy to prove that many factors can constrain and bias the judgment, but this only means that we should get rid of such factors or at least try to counterbalance their effects.

It is always tempting, of course, to take as valid the measure that is the most reliable, but a measure can sometimes be reliable simply because it is biased by constraints of one sort or another. In short, we can never fully escape the uncertain task of deciding, without external criteria, that a measure does or does not assess the thing we are interested in. Fortunately, this judgment does not belong to any one scientist. In the long run, it is the scientific community that will decide the issue.

SUMMARY

These studies undertake to develop and refine a method for the quantitative estimation of sensory magnitudes in the field of audition. One form of the method of *magnitude-estimation* utilizes a standard stimulus and a set of variable stimuli. The standard is assigned some convenient modulus, *i.e.* its loudness is assigned a number such as 1, 10, or 100, and O's task is to assign numbers to the variables in a manner that reflects the magnitude of the ratio between standard and variable. Several different experiments conducted by this procedure gave consistent results over wide ranges—ranges as great as 1000 to 1 in subjective ratio (corresponding to a physical ratio of 90 db.).

Another form of this method dispenses with the standard, and O merely assigns whatever numbers seem appropriate to describe the loudness of a series of intensities presented in irregular order.

The median magnitude estimations obtained with both procedures were consistent with a loudness-scale that is a power function of physical intensity: $L = kI^{0.3}$. This scale has been derived from the pooled results of about a dozen different studies carried out by other investigators.

The assets and liabilities of the method of magnitude estimation are illustrated and discussed.

THE OVERLAPPING OF SIGNALS FOR DECISIONS

By J. F. MACKWORTH and N. H. MACKWORTH, Cambridge, England

Nearly a century has now elapsed since the first demonstrations of the span of apprehension. Better experimental techniques later became available, but in 1926 when Gill and Dallenbach reviewed the literature, they emphasized that not all the questions being studied under the heading of span of apprehension or range of attention were entirely meaningful.[1]

Similar warnings have recently come from Vernon who points to dangers in an approach limited to studies of the estimation of the numbers of objects present, *i.e.* without regard to the groupings of these separate elements into an easily apprehended structure.[2] Graham also currently mentions the great need for an adequate theory of grouping and relatedness.[3] Another valuable survey of the present position of research on the span of apprehension is given by Woodworth and Schlosberg.[4] The number of simultaneously seen objects is stated as being about eight when these are black dots scattered on a white background at an exposure of 0.10 sec. The effect of grouping is considerable. In fact as many as 25 dots can usually be seen if these are presented as 5 groups of 5 dots each arranged as on a playing card.

A landmark in the experimental literature was the demonstration by Glanville and Dallenbach of the need to consider the marked effect on the range of attention of the cognitive requirements of the task.[5] The greater the demands on cognition the less the range of apprehension. For example, counting geometrical forms gave a value (50% correct) for the span of apprehension which was double that obtained from the same Os when asked to name the forms or to name and to give the color of the forms.

The importance of previous perceptual experience was recognized earlier. Tachistoscopic studies by Erdmann and Dodge proved that four or five times as many capital letters were correctly reported when these were presented as familiar words rather than as disconnected items.[6] Recently Gibson and Gagné have shown that training can improve performance at reading digits or at estimating the number of

[1] N. F. Gill and K. M. Dallenbach, A preliminary study of the range of attention, this JOURNAL, 37, 1926, 247-249.
[2] M. D. Vernon, *A Further Study of Visual Perception,* 1954, 55.
[3] C. H. Graham, Visual perception, in S. S. Stevens (ed.), *Handbook of Experimental Psychology,* 1951, 902.
[4] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology,* 1954, 72-106.
[5] A. D. Glanville and K. M. Dallenbach, The range of attention, this JOURNAL, 41, 1929, 207-236.
[6] Benno Erdmann and Raymond Dodge, *Psychologische Untersuchungen über das Lesen,* 1898, 140.

counters present in a display.[7] They doubt, however, whether such training improves the general efficiency of perception. There were no signs of any difference between the trained and untrained groups in their scores on a flexibility of attention test. S had to scan five schematic instrument dials projected on a motion picture screen in order to report any abnormal readings from the five pointers moving continuously.

Special problems on the distribution of awareness are found when S has to include two or more stimulus-fields at the same time—particularly when these are difficult to combine. Hylan studied this 50 yr. ago in a task in which S had to total the number of lines in one, two, three or four series exposed simultaneously through slits by revolving drums. There was a marked slowing when the extra series meant that S had to deal with a very irregular sequence. The extra series could, however, speed performance when there was a predictable regularity in the final sequence. Hylan states that "the time required to count a single line was 0.44 sec. for the single series, 0.31 sec. for the double series, 0.28 sec. for the triple series, and 1.02 sec. for the quadruple series. . . . In the single and double series the order in which the lines were shown presented no difficulty in being formed into a rhythm that could be remembered. This was also true of the triple series if the order were not too complicated. A triple series made too irregular to allow being remembered was used which required 0.48 sec. or nearly twice the time required for the simpler triple series already reported. . . . The quadruple series was also too complicated to be remembered."[8]

Vernon quotes 0.2–0.3 sec. as the time needed for S to shift his attention from one to another of two quite unrelated successive events.[9] Broadbent found that alternation of attention between two voices saying pairs of random digits can be successfully undertaken when each pair arrived every 2.0 sec.[10] This period included the time needed for two perceptions and two shifts of attention. Poulton has demonstrated that when two stimuli were separated by a brief interval, the difficulties in shifting awareness from the first response to the second stimulus depended on a lack of readiness in S, not to his having to recover from what he has just done but to his having to prepare himself for what he has to do next.[11]

Conrad has emphasized that most studies on the distribution of awareness have considered concurrent tasks which were quite unrelated.[12] With investigations of a multiple dial display he found that the number of omissions per minute (N) depended on the number of signals per minute (S) in a way expressed by the equation: $\log N = a + b \log S$. Studies of the moments at which responses were made also brought out that more dials increased the timing error. For any given speed, doubling the number of dials from two to four approximately doubled the

[7] J. J. Gibson and R. M. Gagné, *Motion Picture Testing and Research,* 1947, 134-136.

[8] J. P. Hylan, The distribution of attention, *Psychol. Rev.,* 10, 1903, 380.

[9] Vernon, *op. cit.,* 212 f.

[10] D. E. Broadbent, The role of auditory localisation in attention and memory span, *J. exp. Psychol.,* 47, 1954, 194-195.

[11] E. C. Poulton, Perceptual anticipation and reaction time, *Quart. J. exp. Psychol.,* 2, 1950, 111.

[12] R. Conrad, Speed and load stress in a sensori-motor skill, *Brit. J. indust. Med.,* 8, 1951, 1 f.

size of this form of error. It was interesting that changes in speed did not affect this alternative criterion of performance. Conrad has also shown with a similar task that increasing the number of dials from 4 to 12 increased the number of occasions per unit time on which pointers were permitted to stop: this change also increased the mean duration of such events.[12]

Objectives. The primary aim of the present paper is to draw attention to certain physical characteristics inherent in visual situations which are simultaneously presenting rapidly changing events overlapping in time. Secondly, an attempt has been made to measure varying stimulus-situations of this kind more exactly, to specify the changes that occur from moment to moment during a simple laboratory task. Thirdly, the meaningfulness of the proposed index of task-difficulty has been assessed by considering the relationship between momentary changes in this index and fluctuations in average achievement at this work which demands a series of decisions.

Briefly, signal-overlap will account for most of the psychological effects of simultaneous visual presentations—but no claim is made that this is more than one chapter in the story. Further possible factors are now being experimentally evaluated.

Speed in decision-taking can be a relatively simple concept under *single-channel* conditions. Here a stream of separate problems appears through a single physical gateway, one problem at a time. A new problem arrives every few seconds at approximately regular intervals. The relevant facts appear abruptly and cannot be predicted or considered in any way beforehand. Action demanded by the events must be taken instantly, before the stimulus-objects disappear, as the information cannot be remembered. The serial order in which the decisions are to be taken is predetermined by the situation. Single-channel speed-stress is easy to measure because the stimulus-situation can obviously best be specified in the ordinary way by taking the average rate at which the environment is demanding decisions. The successive events in the single channel are more or less evenly spaced along the time-scale, and therefore this average speed at which decisions are demanded is a good index of the difficulty of the task.

With *multichannel-conditions* the situation is rather different; here the average speed is quite inadequate as a measure of stress and some other index is needed. A multichannel display has two or more signal sources, each of which is providing a stream of signals. The data appear through more than one physical gateway, hence the demands for immediate action will overlap in time.

An extra channel added to a display without extra time will obviously cause trouble because that channel calls for an increased number of decisions per minute. The main thesis of this paper is that multichannel-displays are undesirable for a less obvious reason. Multichannel displays create difficulties even when there is no change in the average rate at which decisions are demanded, *i.e.* even when time is added directly in proportion to the increased number of channels. For example, if the number of channels is doubled, it is not enough to halve the number of

[13] Conrad, Some effects on performance of changes in perceptual load, *J. exp. Psychol.*, 49, 1955.

signals that occur over a given period of time in each separate channel. Conrad has also shown that achievement in multichannel displays cannot be specified by quoting simply the average speed required from S without mentioning the number of sources of signal he has to consider.[14]

An arrangement whereby more gateways supply the same number of signals within a given period will tend to give the effect that action-demands are irregularly spaced along the time-scale. There are greater opportunities for a large number of signals to occur at once. The speed-level varies much more from moment to moment and there is therefore a need for a measure which will estimate rapid changes in speed-stress, to assess the speed-stress under which each separate signal is considered.

Cox and Smith analyzed the problem from a general statistical point of view. They considered the matter largely as a question of the frequency-distribution of the time-intervals between successive events to be expected when several strictly periodic sequences of events are superimposed on one another to form one final common stream.[15] The variables considered are the number of sources, the relative signal frequency of each source and the average signal frequency of their pooled output. In a later paper Cox and Smith have taken up the problem of the pooled output from several independent sources when the intervals between successive events at any one source are assumed to be independent random variables all with the same distribution. The greater the number of original sources, the closer will the pooled distribution be to an exponential distribution; more very short and very long intervals are found.[16] Conrad has considered these statistical principles in relation to performance at a sensorimotor task.[17]

Problem. The problem is to measure and express temporary bursts of intense activity on the display; this requirement becomes especially noticeable in decision-taking when this would be within the limits of ability if the problems were more evenly spread over the available time. The difficulty of a given signal in a multichannel display cannot be assessed only in terms of the time-intervals between pairs of successive signals; groupings may occur in which the perception of each signal is influenced by several other signals, as well as by those nearest to it in time. This is particularly so in taking decisions each of which has to be considered over a definite period—for at least several seconds—rather than for a single fleeting instant.

It seems reasonable to estimate the momentary difficulties arising from these close groupings of events by the extent to which any given signal overlaps in time with any other signal in the series. This index of overlap for a given signal is therefore suggested for the assessment of these vitally important and crowded moments in the task, when everything seems to happen at once.

[14] Conrad, *Brit. J. indust. Med., op. cit.*, 4 f.
[15] D. R. Cox and W. L. Smith, The superposition of several strictly periodic sequences of events, *Biometrika*, 40, 1953, 1.
[16] Cox and Smith, On the superposition of renewal processes, *ibid.*, 41, 1954, 97.
[17] Conrad, Missed signals in a sensori-motor skill, *J. exp. Psychol.*, 48, 1954, 7-9.

METHOD AND PROCEDURE

The procedure adopted in applying this technique is best indicated by considering its actual use in the analysis of a motion picture display of the type to be described later. Each problem was visible for 14 sec. and this active period was marked out for every signal in the series along a scale accurate to the nearest second. Separate horizontal lines were allocated to each of the channels in the display, but all the channels shared the same common time-base. The total length of time that each signal was overlapped by any other signal could therefore readily be measured from this chart. The *index of overlap* for a given signal was the total sum in seconds of the various periods during which that signal was overlapped by any of the other signals in the series.

This may be illustrated by considering a series of three signals, A, B, and C, each visible for 14 sec. (see Fig. 1). If A overlaps the first 12 sec. of B and A
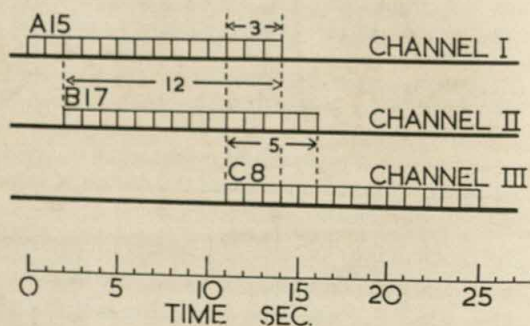
FIG. 1. INDEX OF SIGNAL-OVERLAP

also overlaps the first 3 sec. of C, then the signal overlap-index for A = 15 sec. Similarly, B will overlap 12 sec. with A and 5 sec. with C; therefore the signal overlap index for B = 17 sec. As C will overlap 3 sec. with A and 5 sec with B, the signal overlap-index for C = 8 sec.

Fig. 2 contrasts the physical situations that result when this measure of signal-overlap has been applied to a 4-channel and a 12-channel display. In either case 50 signals have arrived in 5 min. The average speed-stress is, therefore, exactly the same—an average of one problem every 6 sec. There are, however, obvious differences between the two forms of the task, despite the exposure-time for every signal being 14 sec. in both examples. More channels mean more chances of overlap. When each of these channels is firing at approximately regular intervals, but at a rate different from that of any other channel, action-demands are more irregularly spaced along the time-scale. This irregularity is shown in terms of time in Fig. 2 since the index of overlap for each signal is entered on the scale at the moment of onset of that particular signal. The effect of this same irregularity is more readily seen from Fig. 2 in terms of the marked peaks in the overlap-readings found with the 12-channel display but not with the 4-channel presentation.

These peaks are usually separated by particularly marked spells of enforced in-

activity. The lulls are of no benefit to performance in some forms of task, al-
though they do, of course, keep down the average speed at which work is de-
manded. It is interesting that these lulls are also misleading in another way be-
cause they partially disguise the extent of these vicious upward swings in the read-
ings of overlap. An analysis of this kind seems necessary for a fuller understanding
of the situation; for example, the average signal-overlap for the 12-channel display
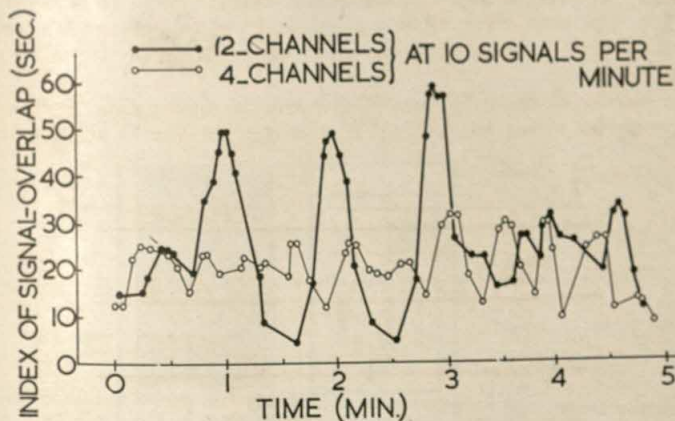
FIG. 2. MULTICHANNEL DISPLAY SIGNAL AND OVERLAP

is not very much higher than that for the 4-channel—29.2 sec. compared with 20.5
sec.

Method. An experimental task was devised to test the general hypothesis that
peak demands like these were particularly harmful to accuracy. This equipment
gave a series of problems in which S had to match two sets of facts directly and
immediately available in his environment. Essentially this was a task on the match-
ing of cards; one member of each pair of cards was always the same and remained
constantly on view through a window. Other cards were brought one at a time
alongside this fixed card by means of a slowly moving belt. This provided a series
of comparisons in which S said how many symbols were common to the fixed and
moving cards on each occasion. (To make the task more meaningful for S, this
was presented as a comparison between two simplified 'flight-plans.' Each card
had six items of information on a given aircraft and S was therefore supposed to
be reporting the extent of the similarity between the information received from two
different pilots.)

Apparatus. Fig. 3 shows one such window at a stage when a moving card has
newly arrived for matching. Here the correct answer would be 'B, Two.'

On different test-runs, there were between 2 and 12 such windows in the dis-
play. Although several windows could simultaneously demand attention, all the
windows in a display were not necessarily showing pairs of cards at the same
moment. The various belts moved at slightly different speeds—the ratio of the

fastest to the slowest was never more than 2:1—hence the pattern of visible cards was continually changing, but S usually had no very great difficulty in deciding the order in which he should take the cards. The complete display covered an area 14 × 14 in. and was set at right-angles to the line of sight. The mid-point of the display was about 36 in. from the eyes. It was therefore unnecessary for S to move his head since the display could be read by making eye movements of not more than ± 11° either laterally or vertically.

Procedure: (1) Initial experiment. Eighteen volunteer Naval enlisted men be-
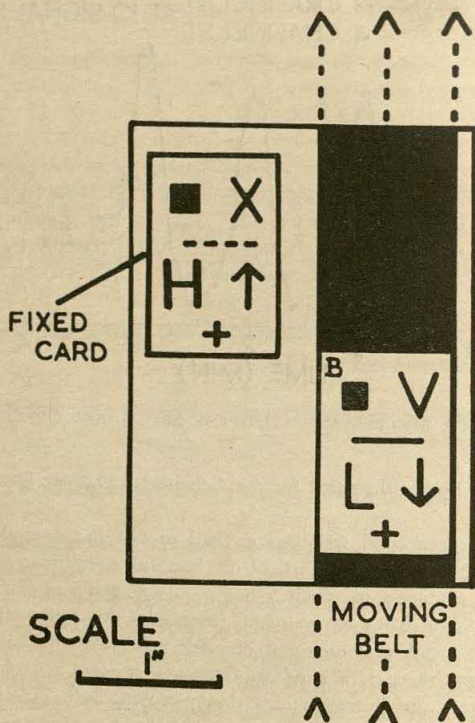
FIG. 3. EXAMPLE OF A SINGLE-CHANNEL DISPLAY

tween the ages of 18 and 31 yr. took part in the first investigation. They were tested singly on a series of 12 runs with this multichannel display. Six runs were given in the morning after a practice run, and 6 in the afternoon with a further practice run. Each of these runs lasted several minutes and consisted of 100 comparisons between pairs of cards. Two factors were varied; speed and number of channels.

The *speed* was set at one of six levels which determined the average number of seconds allowed S for each comparison. On different runs, speeds were selected which represented for most people the range between a leisurely pace and extremely

fast work. These six speeds of 10, 9, 8, 7, 6 and 5 sec. per comparison meant that the whole moving card was visible for about 15, 13.5, 12, 10.5, 9 and 7.5 sec. respectively.

The *number of channels* was altered by changing the number of windows for different runs of the test. One of six displays was selected to expose 100 moving cards in the given standard time through either 2, 4, 6, 8, 10 or 12 windows.

Each S did each of the 6 speeds on 2 occasions and Latin-square planning ensured that the 12 runs for each S included 2 runs from each of the 6 displays.

The exposure-time was kept constant for a given speed; the actual size of the display windows was reduced whenever there was any increase in the number of channels. The reason for this was that the belts carrying the cards had to move more slowly when there were more windows on the display to ensure that the same number of events occurred in a 5-min. run. The implication is that any difficulties in performance found with more display channels could not have been directly due to changes in the exposure-times. On the other hand, changes in speed did alter the exposure-times and these could have affected performance.

No window ever showed more than one moving card at any stage in the presentation. Another point is that each window became active at approximately regular intervals but the frequency of events was different for the different windows.

(2) Serial experiment. Twenty further Naval enlisted men aged 18 to 28 yr. took part in a more detailed consideration of performance at this task. Particular attention was now paid to a temporal analysis of the effects of presenting the signals through different numbers of channels.

An attempt was made to hold the average speed constant and as close as possible to 6 sec. per pair of cards, the whole of every moving card being exposed for about 9 sec. The effective exposure-time was, however, 14 sec., from the moment the upper half of the card appeared until the lower half started to disappear. Each run lasted 5 min., during which about 50 events occurred. Five different displays showed these events through either 4, 6, 8, 10 or 12 windows. One important change in the procedure was that the test-material was presented by means of a 35-mm. motion picture film to ensure that the immediate time-relationships of overlapping signals were held constant. Although the dimensions of the display remained as before, there was inevitably some loss of definition in the projected image.

Every S undertook a series of 10 runs, all 5 displays being presented once in the morning and again in the afternoon. The films were given in a Latin-square arrangement, varied for each group of 5 Ss. The morning and afternoon sessions each took little over an hour and both sessions were preceded by a practice film.

RESULTS

(1) Initial experiment. The results are given in Table I and summarized in Fig. 4. The data on the influence of the number of channels have been grouped under the headings of *few channels* and *many channels,* for 2–6 and 8–12 windows respectively. The criterion of performance was the percentage of failures in the comparisons. This measure was the percentage

of wrong plus the percentage of missing answers. This was an average for a group of Ss, since there were 6 readings for each of the 36 experimental conditions.

(a) Average speed-stress. Statistical analysis confirmed the presence of a harmful effect on performance due to the average demanded rate of working. At the speed-stress level of 6 sec. per comparison there was a statis-

TABLE I

PERCENTAGE INCIDENCE OF FAILURES

No. channels		Seconds per comparison					
		10	9	8	7	6	5
Few	2	3.3	4.0	8.7	4.3	10.3	29.7
(2–6)	4	4.3	5.0	7.3	8.3	12.7	30.7
	6	3.7	8.3	8.0	7.7	13.0	19.0
Many	8	10.0	11.6	14.7	18.0	23.3	29.7
(8–12)	10	11.6	10.0	8.3	14.7	23.0	20.3
	12	9.3	9.3	12.7	12.3	22.3	27.7

tically reliable increase in failures compared with scores at the 10-sec. speed. This decrement due to an increase in the average speed-stress was found with the few-window and the many-window situation. In the former instance the average percentage of failures rose from 3.8% to 12.0%, in the latter it rose from 10.3% to 22.9% failures. Reliable differences were found on testing the significance of these percentage-differences, since the ratios of the mean differences to their standard errors were 5.4 and 8.4 respectively ($p = <0.003$).[18]

(b) Number of channels. The effect of increasing the number of channels is clear from Table I and Fig. 4. Performance was worse with many windows than with few windows for a given average speed. This decrement due to more channels was considerable; doubling the number of channels usually meant about twice as many failures, even when the average speed was kept constant.

This multichannel effect was present at nearly all levels of speed-stress, see Fig. 4. At the *slow* working speeds, few windows gave 4.8% failures and many windows 10.3%. Similarly, at the *medium* speeds few windows averaged 7.4% failures and many windows 13.4%. At the *fast* speeds 19.2% failures occurred with few windows and 24.4% with many. A statistical analysis established the reliability of these differences and the ratios of the mean differences to their standards errors were 9.0, 8.3 and 5.2 respectively, $p = <0.003$. The very highest speed of all was obviously,

---

[18] G. U. Yule and M. G. Kendall, *Introduction to the Theory of Statistics,* 1940, 352.

however, an exception, see Fig. 4. Here, in fact, there was a dramatic rise in failures which seemed to be just as marked with few windows as with many.

(2) Serial experiment. The first question was to determine the most
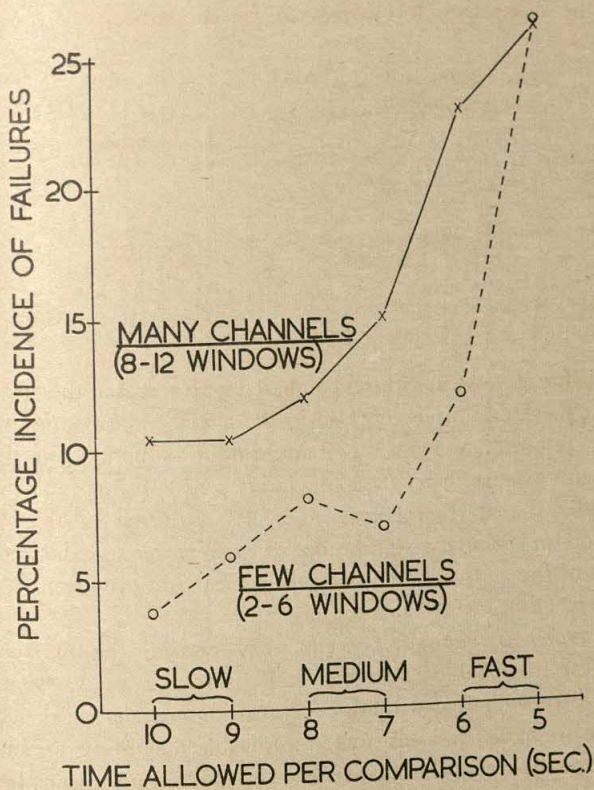
FIG. 4. RELATIONSHIP BETWEEN FAILURE AND SPEED FOR MULTICHANNEL AND FEW CHANNEL DISPLAYS

Multichannel displays give more failures for a given level of speed.

suitable measure to apply to performance at the task. In the previous experiment the analysis was based on total failures, the sum of the wrong and the missing decisions, but the present data suggested that a better index might well be the omission-scores by themselves. Detailed comparisons therefore follow between signal-overlap and omission-scores, but this evidence will be introduced by a rather more general consideration of the results.

The analysis turned on the extent of the relationship between the index of overlap for a given card and the shortcomings in performance on that card as is revealed by the failures, *i.e.* any wrong or missing answers.

Fig. 5, for example, illustrates the typical trends for the 5-min. run taken from the 8-channel display. The card-comparisons during this period
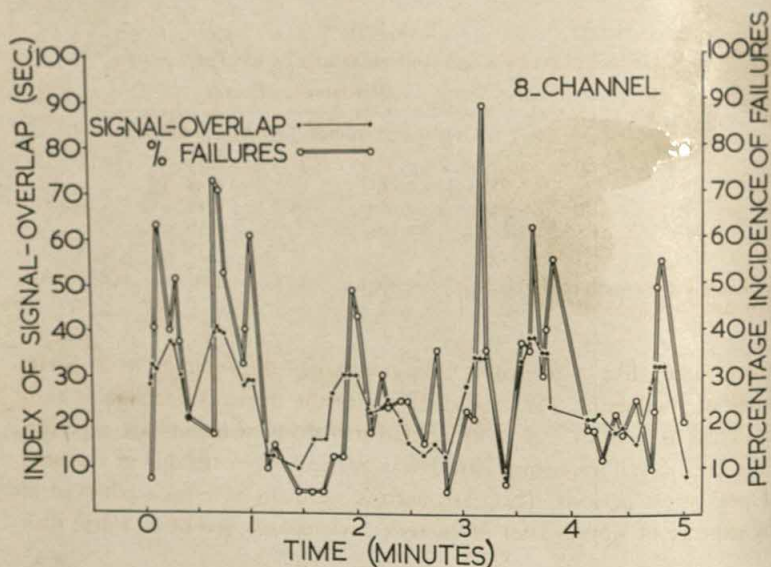
FIG. 5. RELATIONSHIP BETWEEN INDEX OF SIGNAL-OVERLAP AND PERCENTAGE INCIDENCE OF FAILURES

have each been represented there by two readings—the index of overlap and the percentage of failures for that card. (Each card was read 40 times, *i.e.* twice by 20 Ss, and the incidence of failures calculated on this basis.) Inspection suggested some similarity between estimated and actual difficulty from moment to moment during the run. Further statistical analysis confirmed this impression; Table II indicates that for the 8-channel display the correlation coefficient between overlap and failures was r = 0.66.

All the correlation coefficients in Table II showed very definite signs of association between this index of overlap and faulty performance (p = <0.001; this was the case with either failures or omissions as the measure of performance).

A study of Fig. 5 (and similar graphs from the other four displays) prompted the thought that these correlations mostly arose from an associa-

tion between the estimated and actual peaks of stress, rather than at the more humdrum levels of activity. This finding is in line with the tendency to be seen in Table II for the relationship to become even closer when the number of channels was increased. This is believed to have arisen because more channels in a display will magnify the peaks in the demanded

TABLE II

RELATIONSHIP BETWEEN SIGNAL-OVERLAP AND FAULTY PERFORMANCE

No. of channels	Correlation-coefficients	
	Overlap and failures	Overlap and omissions
4	0.53	0.48
6	0.66*	0.52
8	0.66	0.58
10	0.68	0.56
12	0.72	0.73

* Excluding one-sixth of the readings, those from a badly focussed window; r = 0.35 when these data were included.

stress, rather like a telescopic lens emphasises the summits in a distant range of mountains. Fig. 5 also illustrates the tendency for signals early in a crisis to be dealt with more satisfactorily than one might have expected. This lag of failure behind stress was particularly noticeable at the start of peak-stress periods. This early success seems to have been achieved at the expense of signals later in the series, when work was often worse than

TABLE III

SUMMARY OF AVERAGED DISPLAY AND PERFORMANCE

No. of channels	Conditions		Results	
	Signal-overlap (sec.)	Speed (sec./card)	% failures	% omissions
4	16.5	7.1	11.7	4.8
6	20.6	6.3	21.7	11.8
8	25.5	5.7	30.6	19.9
10	23.2	6.8	25.5	17.9
12	25.9	5.7	30.4	23.0

one might have predicted. A slight and temporary rise in wrong answers was usually followed by this very high proportion of missed decisions. Recovery was, however, usually extremely rapid.

Table III relates the average failure and omission-scores to the average physical readings for each of the displays, *i.e.* to the speed and the index of overlap.

These general readings indicate a relationship between number of channels and overlap on the one hand, and performance on the other. Table III also shows that the speed factor has complicated the issue because it was not as constant as it should have been for all the displays. The results cannot, however, be attributed simply to these unexpected variations in speed. For example, the 10-channel display was done badly compared with the 6-channel, but the 10-channel display was in fact slightly slower. The index of overlap does take into account to some extent such changes in the speed, *e.g.* the index of overlap will rise if rather more signals (each of the same length as before) are crowded into the same period of time.

(a) *Percentage of omissions and the index of overlap.*

Table III emphasizes that about one-half to three-quarters of all failures arose from omissions; this tendency was even more marked when sudden

TABLE IV

Regression-Line Data

Grand means

Source	$(\bar{x})$	$(\bar{y})$	Regression coefficients (b)*
	Signal overlap (sec.)	Omissions (%)	
Grouped means	28.0–34.9	26.4	1.3
Ungrouped data	22.6	16.0	1.5
Channel-means	15.5	22.3	1.9

* b in the regression-line formula $(\bar{y}-y)=b(\bar{x}-x)$ when y=% omissions and x=signal-overlap in seconds.

extreme difficulties gave severe breakdowns because then omissions usually averaged as much as 80–90% of the failures. Omissions were therefore considered in further detail as a measure of performance. (The number of wrong answers per card was small and usually fairly constant at 5–10%, except when a peak was developing in the index of overlap.)

The first step was to prepare a general distribution-table pooling all the results in this serial experiment. Every card from the whole investigation was entered on this table, the entry being made according to the level of overlap. The percentage of omissions was calculated for each of these levels of signal-overlap. Fig. 6 shows this average incidence of omissions at the eight different levels of signal-overlap with the regression-line fitted to these observed data.

Table IV shows that the regression-line fitted to the same readings without any such preliminary grouping was slightly steeper. As there were un-

equal numbers of cards at the various levels of the index of overlap, this ungrouped line gave a better estimate of the general effect. For every 10-sec. increase in the index of overlap there was a 15% rise in the average incidence of omissions. This regression-line cut the abscissa at 12-sec. overlap;
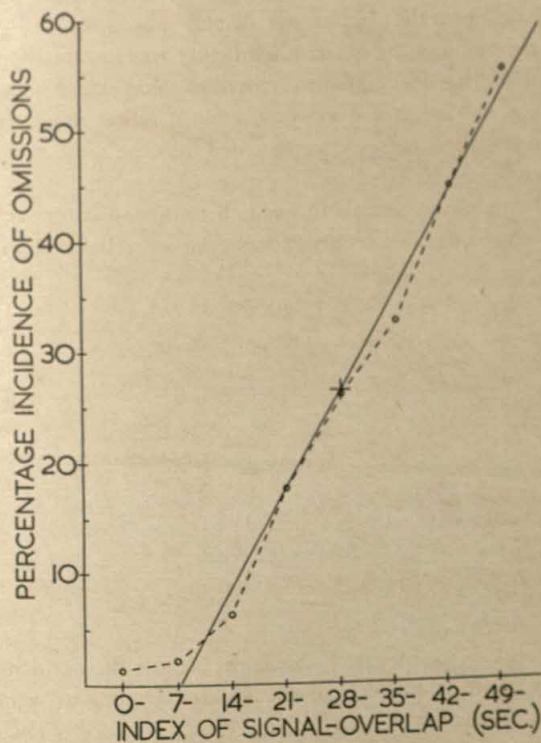
FIG. 6. RELATION BETWEEN SIGNAL-OVERLAP AND INCIDENCE OF OMISSIONS
The higher the index of signal-overlap, the greater the incidence of omissions.

this result meant that two pairs of cards could be read without omissions even when they appeared almost simultaneously.

There was obviously a linear relationship between overlap and omissions. The linearity of the regression of omissions on overlap was highly reliable compared with the departures from the regression ($p = <0.001$, and the analysis of variance for the total data is contained in Table V).

There is an indication in Table IV that further analysis was needed on the data considered separately for each display. The displays with more

channels, of course, had a higher signal overlap score and therefore more omissions, but this rise in omissions with more channels seemed rather faster than it should have been. Table IV shows, for example, that the regression-line drawn through the means for the channels estimated a 19% rise in omissions for every 10 sec. of overlap instead of the 15% rise previously found with the ungrouped data.

Analysis of covariance determined whether there were any differences

TABLE V
LINEARITY OF REGRESSION

Source	df.	sos	ms	VR	Significance
Regression	1	8446	8446	710.6	<0.001
Departures	252	3247	12.9	—	—
Total	253	11693			

TABLE VI
ANALYSIS OF COVARIANCE

Source	df:	sos_y	sos_x	sop	ss_y^1	ms_y^1	VR	Significance
Channels	4	2500	2883	2156	955	239	26.5	0.001
Error	249	9169	20269	11826	2269	9	—	—
Total	253	11669	23152	13982	3224			

TABLE VII
AVERAGE PERCENTAGE INCIDENCE OF OMISSIONS
(Means adjusted for a standard level of signal-overlap.)
Number of channels

4	6	8	10	12
10.8	11.8	12.6	13.6	14.6

between the means for the channels which were not accounted for by this regression-line for the ungrouped data.[19] This was the statistical equivalent of checking whether the omission-scores were all the same had the five displays been compared experimentally at the same level of signal overlap (22.6 sec.).

This indicated that the means were still statistically different from each other in spite of the adjustment for signal-overlap. The nature of the difference is shown in Table VII and Fig. 7.

The differences between means were much reduced by this adjustment, but even so the displays with fewer channels had fewer omissions for the same level of signal overlap. (The adjusted mean for the 4-channel display was significantly lower than that for the 10-channel—$p = <0.05$—or for

[19] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 317-336.

the 12-channel —$p = <0.01$. The adjusted mean for the 6-channel display was also significantly less than that for the 12-channel—$p = <0.05$.)

It was interesting to find from Table VIII that these adjusted means lay along a straight line ($p = <0.001$). Despite the experimental flaws
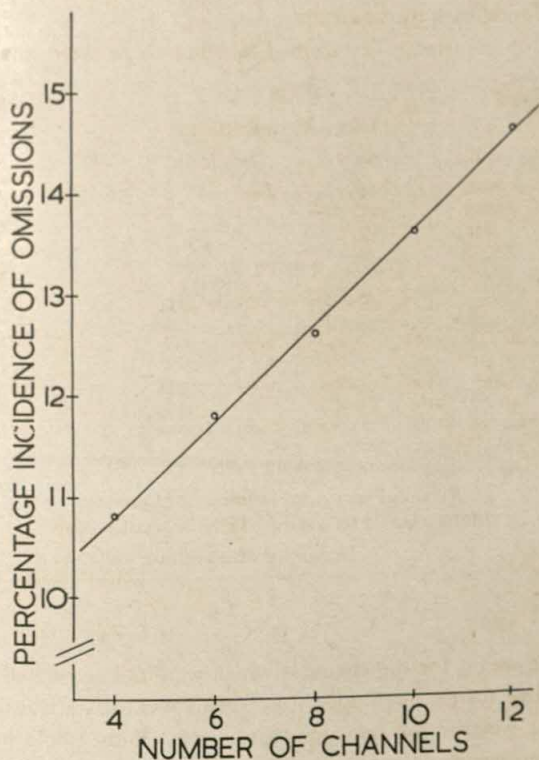
FIG. 7. RELATIONSHIP BETWEEN NUMBER OF CHANNELS AND PERCENTAGE OF OMISSIONS
Multichannel displays lead to more omissions even at a given level of signal-overlap.

previously mentioned whereby uncontrolled variations in the speed-stress factor partially obscured the effects of increasing the number of channels, the correction for signal-overlap removed the reversal noted in Table III and these adjusted means now formed a very regular ascending series. The grand means were $\bar{x} = 8.0$ channels, and $\bar{y} = 15.8\%$ omissions, the regression coefficient (b) being 0.24.

(b) Failures and the index of overlap. Calculations of the same kind were also made to relate the scores on failures to the readings of signal-

overlap. Again a straight line relationship was found for the ungrouped data. The analysis of covariance confirmed that signal-overlap was not the only factor present. The adjusted mean for the 4-channel display was significantly different from that for the 6- and 10-channel at the $p = < 0.05$ level, and from those for the 8- and 12-channel at the $p = < 0.01$ level.

(c) Omissions and the 9-sec. index of overlap. The 9-sec. index of overlap was based on the time that the *whole* of the moving card was

TABLE VIII
LINEARITY OF REGRESSION
(Adjusted means.)

Source	df.	sos	ms	VR	Significance
Regression	1	2.228	2.228	1114	<0.001
Departures	3	0.006	0.002	—	—
Total	4	2.234			

visible. Very similar results were obtained when the data were analyzed again in this alternative way—both for the straight-line relationship between signal-overlap and omissions—and also by an analysis of covariance.

(d) Individual differences. There was a very considerable range of ability between the omission-scores made by the Naval enlisted men who took part in the serial investigation. The best *S* missed only 5% whereas the worst omitted as many as 26% of all the signals. These differences were

TABLE IX
AVERAGE INCIDENCE OF OMISSIONS
(Percentages)

Time	Run 1	Run 2	Run 3	Run 4	Run 5
Morning	20.3	18.4	17.6	17.1	15.3
Afternoon	16.7	16.0	15.2	14.2	14.4

to some extent related to the scores from a written test of intelligence. The coefficient of correlation between omissions and achievement on A.H.4., a test of this type, was -0.51 ($p = < 0.05$).

It was also found that the time required for a comparison of a pair of cards ranged from 6 sec. for the slowest to 3 sec. for the fastest. There was a suggestion in Table IX of a slight effect of practice in the successive omission-scores—but these large differences between *S*s prevented this trend from being statistically reliable.

The reliability of the omission-means for the various displays was

satisfactory. As will be seen in Table X, the average scores for the 20 Ss remained fairly constant when they were re-tested on the same situation.

TABLE X

Average Incidence of Omissions

(Percentages)

Time	Number of channels				
	4	6	8	10	12
Morning	6.1	12.3	21.5	19.5	24.3
Afternoon	3.5	11.3	18.4	16.4	21.7

DISCUSSION

This paper investigates some of the effects of varying the number of sources of visual information upon performance. It attempts to account for these in terms of the simplest possible expression of the situation—the required rate of work from moment to moment. The underlying assumption is that there must be a fixed maximum to the rate at which a man can think effectively when making comparisons between sets of objects. The idea is that this can matter even when the exposure-time is adequate for each separate comparison, *i.e.* in spite of the objects being on view for a period which would be perfectly ample if this were the only match to be made in that time.

The suggestion is that this upper limit can be exceeded in two quite different ways. The obvious and usual way is when time is short throughout a given task—when the average number of decisions required per minute is in general too high. This form of speed-stress is best expressed by the average number of seconds per signal. This index is particularly suitable when the signals are presented through a single-channel display, and are more or less evenly spaced in time.

Displays with more than one source of signal can, however, give trouble even when such average speed-stress is low. Difficulties arise with many channels chiefly because the spacing in time of their pooled signals is so irregular that a crisis is created every now and again. Given that each source is firing at a different rate, irregularity will inevitably occur in the pooled series even if each channel is firing quite regularly. This bunching together of signals can also be regarded as a form of speed-stress—a transitory burst of activity which leads to a momentary overstepping of the permissible limits for the required rate of decision. Such fluctuating demands call for a new index of speed-stress which will measure the demanded rate of

decision for each successive signal, *i.e.* to identify and mark out any treacherous peaks in the stress.

The time between the given signal and the signal nearest to it in the series is an inadequate measure for this purpose, particularly when each signal lasts for a specific period and is not a momentary event. In fact, the index of overlap between signals gives a better understanding of this ever-changing speed-stress with multichannel displays. This index is the total sum in seconds of the various periods during which the given signal overlaps any of the other signals in the series.

Refinements and modifications are needed to extend the usefulness of this index in analysing behaviour in serial tasks. For example, some adjustment is obviously required to correct for any alteration in average speed-stress, when this is also reducing the time during which the signals are exposed.

Already, however, this index of signal-overlap is demonstrably of some value in understanding the reasons for faulty serial performance, whether this is in terms of failures or omissions. It can be used within a series of signals of equal length occurring on a given number of channels at a given average number of signals per minute.

Quick bursts of speed-stress are the most important effect of adding further sources of signal—provided that there is no general alteration in the average speed at which decisions are demanded. An increase in the number of channels in a display has a further effect on performance, the reasons for which remain as yet obscure. This additional disadvantage of too many channels appears to be directly proportional to the number of channels.

Further experiments are needed to understand this interesting multichannel effect which is *not* due to speed-stress as measured by the extent of overlap between signals. Although the present investigations were not designed for this purpose, some speculation is perhaps permissible on the possible causes of these residual effects. It may be that the index could be weighted to correct for this by making some allowance for the *number* of signals involved in a given overlap. This could be tried in practice, but the index would no longer be strictly speaking a straightforward measure of speed-stress intended to elucidate these environmental effects.

Further studies, especially with other related tasks, will determine whether this multichannel effect apart from signal-overlap was real or an artifact of the display. The writers believe, however, that several factors could have caused the larger incidence of omissions related to more sources

even when there was the same overlap between signals (see Fig. 7).

In decreasing order of likelihood, the explanations needing experimental investigation are as follows: (1) more channels in a display led to a greater redundancy in the presentation; each window had its fixed card always on view and obviously not all of these would be relevant at any given moment. This might have given rise to some difficulty because there were more signals to neglect in scanning for new arrivals. (2) More channels meant that rather more cards arrived nearly simultaneously for comparison during periods of peak-stress. This could have increased any difficulty in deciding the most effective order in which to take the various comparisons. (3) More channels meant higher and longer lasting peaks of stress. Performance might have worsened because there were fewer of those signals previously described which were done better than expected because they came early in a growing peak of stress. (4) More channels may have caused trouble because the additional irrelevant material required extra eye-movements which now began to be a limiting factor.

One line for further work might be studies of the serial matching of objects on similar but more complex material. Achievement at the taking of decisions would be assessed with special regard to the part played by eye-movements in selecting a series of essential cues from the mass of detail in the immediate visual world. Analysis of the eye-movement patterns in space and time during such tasks may lead to a better understanding of the relationship between the three main psychological aspects of this kind of work. Verplanck has drawn attention to the presence of two of these functions in problems related to the communication of information: (a) the span of apprehension and (b) the immediate memory-span.[20] It is particularly necessary, however, to add a third function: (c) the shifting of awareness, when continuous work is being considered. Vernon has described this as the "successive focussing of awareness upon different parts or aspects of the field each of which can be separately and satisfactorily dealt with, assimilated and comprehended."[21] Grindley found that even with tachistoscopic perception the nature of a peripherally exposed test-object was much less likely to be accurately reported if several other objects were also exposed simultaneously.[22] Often, however, the memory after-image plays an important rôle in the tachistoscopic work with geometric forms and Ss can report from the memory image rather than the im-

[20] W. S. Verplanck, *Human Factors in Undersea Warfare*, 1949, 253.

[21] Vernon, *op. cit.*, 210.

[22] G. C. Grindley, *Psychological Factors in Peripheral Vision*, 1931, 32-33.

mediate percept. The present studies frequently found Ss using their memory after-images of the upper three symbols on one card to match this with their percepts of the corresponding three symbols on the other card. Then the lower three symbols were compared in the same way. Since Ss tended to take three objects at a time, it is particularly interesting to note that Glanville and Dallenbach objectively determined the span of apprehension for geometric forms at about this number of items. They found that different Ss averaged 3.2, 3.9 and 4.3 objects for their span of apprehension when this was taken at the 50%-correct level. These Ss correctly named the forms on most of the cards where only three objects were present, the exposure-time being only 0.08 sec.[23]

Consideration of the rôle of immediate memory in the side-by-side comparison of visual objects could lead to studies of tasks such as the matching of perceived and remembered patterns. Without such a development, these investigations may tend to exaggerate the importance of present perceiving. Alternatively, the experimental situation might be widened to include the coming occasion—the not-quite-yet, to use the term suggested by Bentley.[24] For example, the effects of auditory cueing could be considered especially in the structuring of an ambiguous visual scene. Ways may even be found, in time, to study the overlap between signals for action some of which are imminently foreseen rather than physically present in the environment.

SUMMARY

An attempt has been made to understand more fully the problems facing men in situations in which they have to compare quickly many objects presented visually and simultaneously. Current traditions in psychology may well tend to overemphasize the importance of present perceiving, but experimental studies of relatively complex behavior such as human choice in critical and problematical occasions can perhaps at least start by considering activity in relation to the immediate scene. Human limitations were therefore sought in the rate at which each signal in a continuous series could be accurately matched. The physical measure devised was the index of signal-overlap which is the total sum in seconds of the various periods during which the given signal is overlapped by any other signal.

A task in which objects had to be matched confirmed the great disadvantages for skilled achievement when further physical sources of demands

---

[23] Glanville and Dallenbach, *op. cit.*, 228.
[24] Madison Bentley, Forecast, timing, and other primary factors in the government of certain biomechanical systems, this JOURNAL, 65, 1952, 339.

for action are added to a serial visual presentation. These difficulties are experienced even when there is no change in the average number of signals presented per unit time. The greatest drawback of multichannel displays is believed to be their tendency to give rise to momentary but very damaging peaks of speed-stress. These peaks may pass unnoticed unless the physical situation is analyzed by some measure such as this index of signal-overlap.

There is a definite and rising correlation between this index and faulty performance as the number of channels in the display is increased. A linear relationship has been demonstrated for the regression of missed signals on signal-overlap. The greater the peak-stress, the higher the proportion of failures due to missed rather than wrong decisions.

Most of the multichannel effect can be traced to peak speed-stress. Statistical analysis has however also suggested that more display-channels may have further effects. Experimental work is in progress to determine whether these are generally found or are artifacts specific to this experimental situation. They could have arisen from a greater redundancy in the immediate display, or from an increased difficulty in choosing the most effective order in which to take up the various comparisons.

CIRCLES AND DERIVED FIGURES IN ROTATION

By Hans Wallach, Swarthmore College, Alexander Weisz, Tufts College, and Pauline Austin Adams, University of California

It is a well-known fact that figures change their appearance with a change in their spatial orientation. "Many years ago, E. Mach pointed out that if we turn a square on an edge its appearance differs strikingly from the same objective square when it rests on a side."[1] Orientation alone makes it possible to differentiate between a 'p' and a 'q' in print where they have congruent shapes, and in printing types without serifs the forms of the letters, b, d, p, and q, differ from each other only by virtue of their orientation, all four shapes being congruent. Thus, in the perception of adults, orientation produces properties of form that do not originate in the given shape but that are, nevertheless, indistinguishable from the shape-dependent properties. Spatial orientation of a figure and the resultant form are, of course, dependent on a framework to which this orientation is related. Thus, we may speak of form-properties dependent on framework. In this paper, a different but equally impressive effect of such a framework will be reported.

When a homogeneous circle is drawn or pasted on a background of different color and this pattern is viewed while it is rotated on a turntable at moderate speed, the circle does not seem to turn. This is the case even when the rotation of the background is clearly visible. That under certain conditions a homogeneous circle which rotates about its center may not seem to turn is easily understood. When it is rotated about its center no visible change takes place. This is because any part of the circle can be superimposed on any other part; the circle shares this property with the straight line and the helix. No cues are conveyed to the eye that would indicate whether the homogeneous circle turns or does not turn. Where the surroundings of the circle are visibly stationary, a rotating circle might be expected to appear standing still. As already mentioned, however, the circle sems to stand still even when its background is seen to turn, and this cannot be understood merely from the fact that no sensory data convey the state of motion of the circle. If this were the only factor in the situation the circle would rather be expected to appear to be turning together with its

background. It seems that there is a positive factor operating that causes the perceived circle not to turn. That this is the case is demonstrated by the following observation.

When a circle is placed eccentrically on a turntable, its perceived motion again differs from the objective condition. Its objective motion can be considered as consisting of two concurrent movements. It has a *revolving*, translatory motion, *i.e.* a movement of the circle as a whole around the center of the turntable. It also *rotates* once about its own center during one revolution of the turntable. What happens here is readily seen if one substitutes another figure for the circle. If, for instance, an arrow is
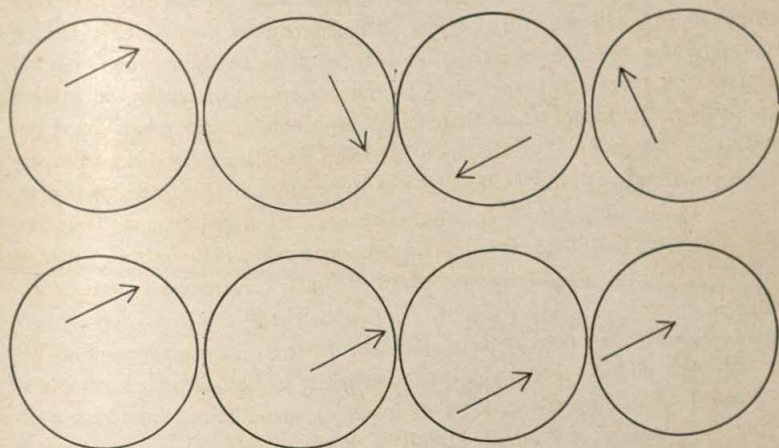
FIG. 1. FOUR POSITIONS OF ARROW REVOLVING CLOCKWISE ON TURNTABLE
Top row: when arrow is attached to table and turns with it.
Bottom row: when arrow is rigged to maintain its direction during rotation.

fastened eccentrically to the turntable, it will point successively in all directions of the compass as the turntable turns through a full rotation (Fig. 1, first row). It thus turns once about itself during such a rotation. This is exactly what the circle does, except that this rotation of the circle is not conveyed to the eye. Its revolving motion, that is, the movement of the circle as a whole about the center of the turntable, is, of course, clearly visible; but the turning of the circle about its center is not seen. As a result the surface of the turntable appears to move in relation to the edge of the circle. To have the arrow move analogous to the experienced motion of the circle, we would have to install a mechanism that would permit the arrow to remain pointed in a constant direction, that is, *not* to rotate,

while it is revolving with the turntable (Fig. 1, second row). In the O's experience, the eccentrically placed circle behaves exactly as the arrow does when it is rigged in this manner. Under these conditions, the arrow would rotate relative to the surface of the turntable, since it maintains a constant orientation while that surface turns about its center. Similarly, the edge of the circle should appear to slide relative to the rotating surface of the turntable, and this is actually observed.

We have pointed out that a homogeneous circle does not convey cues to the eye as to whether it turns about its center. Thus, insofar as conditions of stimulation are concerned, it does not matter whether a circle is fastened to the turntable and turns with it, or whether, like the arrow, it is rigged so as not to rotate when it is going around with the turntable. Briefly expressed: with regard to rotation, the conditions of stimulation are indeterminate. Why, then, is it that except for its revolution, the circle so clearly seems to stand still? If perceptual experience here depended merely on conditions of stimulation, the state of rotation of the circle should remain vague or it should be seen to rotate together with its background. Conditions originating within the perceptual process must be responsible. Seeing the circle stand still as it revolves with the turntable means that its various parts, indistinguishable in shape from one another, remain in some fashion identified. They are identified in relation to the directions of the visual field. The top part of the circle remains top part throughout a revolution, and its left horizontal pole retains the property of being the left horizontal pole, etc. A framework in relation to which the directions of the visual field are defined must be in play.

The framework, however, enters here into a unique function. The framework-dependent properties mentioned in the beginning are a result of 'orientation' of a figure with respect to the framework; the figure has features of shape, identifiable in their own right, that stand in clear relations to the directions of the visual field.[2] Such identifiable features are lacking in the shape of the circle. Instead, identical parts or points are *created* in the perceptual process in dependence on the framework. Identity is imparted to aspects of the circle.[3]

---

[2] Actually the facts are somewhat more complex. See Kurt Koffka, *Principles of Gestalt Psychology,* 1935, 184.

[3] As mentioned above, the straight line shares with the circle the peculiarity that any part of such a curve can be superimposed on any other part. Barring visibility of its ends, no parts or points along a straight line are identifiable on their own. It is interesting that the movement of straight lines also shows peculiarities that imply an imparting of identity to points or parts along that curve, although here the imparted identity derives from other conditions. See Hans Wallach, Ueber visuell wahrgenommene Bewegungsrichtung, *Psychol. Forsch.,* 20, 1935, 325-380.

The notion that the various parts of a circle remain identified as it revolves around the turntable is supported by concrete observations. When two homogeneous rings are so placed on the turntable that they are crossed (Fig. 2) and are turned at a moderate speed, they will first be seen to turn as a unit. Sooner or later, however, they will appear to move independently of each other, each ring having ceased to turn as it revolves with the turntable. In this stage one ring always seems to slide over the other. When one pays attention to one of the points where they cross, one clearly sees one ring move across the other ring. In experience, the ring consists of definite parts which move in relation to the other ring. As stated, these parts originate in the perceptual process being an imparted identity.
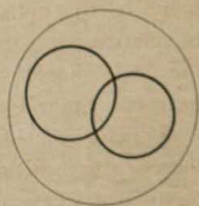
FIG. 2. TWO HOMO-
GENEOUS RINGS ON
TURNTABLE

At this stage, the rings are seen to move exactly as two figures which, by virtue of their shape, have a definite orientation and, like the arrow of the earlier example, are rigged so as to maintain a constant orientation. The triangles of Fig. 3 may serve as an example. As they revolve with the turntable they continuously change their position with respect to each other and cross each other at different points. Whereas they perform this motion owing to a mechanical arrangement, the circles do it owing to our perceptual effect.

An impressive demonstration of the power of this effect can be made with a figure which looks like two incomplete disks attached to each other (Fig. 4). When this figure is first rotated on a turntable it is, as a rule, simply seen to turn. When, however, it is observed for a little while, a change takes place. Two complete disks seem to revolve about each other but seem not to rotate. Where they
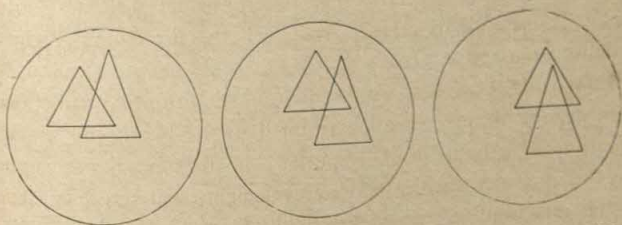
FIG. 3. THREE ARRANGEMENTS OF TWO TRIANGLES ON TURNTABLE
Rotation clockwise

overlap, one of them clearly appears to be on top of the other one and to slide over it. Some Os even 'see' the contour of the upper disk in the region of overlap. (To understand this latter observation it must be kept in mind that the disks appear to change their position with regard to each other as they revolve, such that continuously different parts of a disk move into the region of overlap. Thus, the part of the disk which seems to be on top of the other one at one moment was seen just before bordering on the turntable surface where its contour was seen against a ground of different color.) The whole process is essentially the same as

that observed with the two rings. In the present case, however, our effect also causes two figures to appear instead of one and, for some Os, causes a contour to appear on an objectively homogeneous surface.

Another demonstration of the strong tendency of a circle to appear stationery can be given with a wide ring (about 2 in. wide when its diameter is 8 in.) whose outer edge is smooth and whose inner edge is scalloped. When such a figure is slowly turned about its center, the scalloped edge is, of course, seen to revolve, but the smooth outer edge appears to stand still. While a certain part of the circular outer edge is seen to remain in its place, the part of the inner scalloped edge right across from it seems to move on. This gives rise to the odd impression of a continuous shearing in the surface of the ring. Again, the merely imparted identities of the parts along a circular edge function with all concreteness.

It is interesting that this imparting of identity may also take effect with figures which to some degree depart from circular form and have an orientation by virtue of their shape. A deformed disk whose shape is shown in the oval outline of Fig. 5 may serve as an example. When it is rotated about its center it seems to stand still,
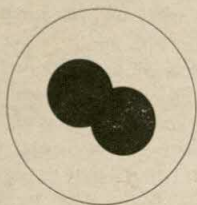
FIG. 4. TWO INCOM-
PLETE DISKS AT-
TACHED TO EACH
OTHER

FIG. 5. A DEFORMED
CIRCLE COVERED BY PAT-
TERN OF RADIAL LINES

although it has a distinct orientation when it is seen at rest. Instead of a rotation of the whole figure, a continuous deformation seems to take place; a bulge is seen moving around and around like a wave. This peculiar experience is itself the result of the imparted identity. The schema according to which identity is imparted can be visualized as a pattern of radial lines that lies over our figure such that its center coincides with the center of the figure. An intersection of a radial line with the contour of the figure represents a point perceived as identical. Such an intersection would remain approximately in place if our figure were turned under the stationary pattern of radial lines, and the same would be true of points perceived as identical on the contour of the turning figure. When the bulging part of the contour passes by, such an identical point first moves outward and then inward just as an intersection between the contour and a radial line would do. Thus, the passing of the bulging part causes motion primarily in the radial direction. The radial motions of the various points add up to the impression of a traveling deformation.

The well-known expanding or contracting of a rotating spiral is another ex-

ample of the effect of imparted identity. An Archimedean spiral of moderate pitch, presented concentrically in the circular aperture of a screen that conceals its external end, will show no turning motion at all when it is rotated about its center; it will only seem to expand or contract. A part of the spiral that lies in a certain direction from its center is seen to remain in that position, while objectively it moves to one side and is replaced by an adjacent section where the curve has, say, a larger radius. This happens simultaneously to all parts. From moment to moment, objective sections having a greater distance from the center come to be located in certain radial directions. And, since it is to these locations that identity is imparted, every part is seen to move in a radial direction and the whole curve is seen to expand.

Deformations displayed by a rotating ellipse that is not much different from a circle can be explained in a similar way. When an ellipse whose axes measure 25 and 23.5 cm. is placed on a contrasting background and rotated about its center, it appears to stand still while its contour seems to pulsate. A perceptually identical part of the contour moves alternately inward and outward in a radial direction.

Deformations much stronger than the ones just reported can be observed when an ellipse is used that departs more from the circular form. They have been described by Musatti,[4] who from such observations went on to investigate the three-dimensional effect that can occur with such ellipses; the latter will be discussed below. We found that the largest effects could be observed with an ellipse in which the lengths of the major and minor axes were approximately in a ratio of 3 to 2. In that case the deformations were so strong that the whole figure appeared to be fluid. Its motion is hard to describe. It is by no means deformation of a stationary form alone; the figure also turns. Probably its deviation from circular shape is so strong that the change in orientation makes itself felt. Identity is also imparted to locations along the ellipse and this produces its deformation. As turning and deformation occur together a peculiar amoeba-like motion results. It should be added that for some of our Os the deformation was somewhat restricted and did not affect the region in the immediate neighborhood of the poles of the major axis. For these Os the figure was perhaps too strongly elliptical. A still narrower ellipse is seen to revolve as a rigid form by all Os.

The motion of the ellipse is the same, no matter whether it is rotated about its center or whether it is placed eccentrically and revolves on the turntable. Fixation does not affect it. For instance, when it is placed eccentrically it does not matter whether the center of the figure or the center of the turntable is fixated. The same holds for observation of circles and other figures. Thus, eye-movements are presumably not the cause of our effects. As is to be expected no deformations occur when the ellipse revolves on the turntable but does not rotate, that is, when it is eccentrically attached to the turntable and, owing to a suitable mechanism, is kept constant in orientation throughout.

It should also be mentioned that the deformations of the ellipse did not appear immediately when the turntable was set in motion. It took from 5 sec. to 1 min. before a naïve but properly instructed O reported that the figure looked fluid. This delay can be avoided, however, if O is not permitted to see the ellipse at rest. All

[4] C. L. Musatti, Sui fenomeni stereocinetici, *Arch. Ital. Psicol.*, 3, 1924, 105-120.

of 10 naïve *O*s saw the figure immediately as fluid when the turntable was already turning at the start of the observation period. On the average, they reported this 4 sec. after the start of the observation, whereas 10 other *O*s who were allowed to see the figure first at rest reported that the figure had become fluid on the average 11.5 sec. after the turntable was set in motion, a difference reliable at the 5% level of confidence. The reason for the delay under the latter conditions is still obscure.

THE STEREO-KINETIC PHENOMENON

When one observes the deforming motion of the ellipse lying in a plane for sometime with one eye, a further change may occur. The figure appears to rise out of the plane of the turntable as a circular disk. For instance, when a pattern like the one shown in Fig. 6 is slowly turned, an *O* viewing it monocularly may experience two changes in the perceived process, only the first of which has so far been discussed. At the outset he sees a rigid figure revolving about the center of the turntable much like the white ellipse when it is seen at rest. Next the ellipse seems to be fluid and moves around in wallowing amoeba-like motion. Finally, this ever-changing form may turn into a rigid circular disk which rises out of the plane of the turntable at an angle of about 45° and seems to roll around its center in this slanted position. That such a three-dimensional effect can be obtained with a rotating ellipse was apparently first discovered by Musatti who described this effect in the paper previously mentioned. In the same paper Musatti deals with still another depth effect which also emerges when a rotating plane pattern appears to distort. Inasmuch as it occurs with rotating patterns of circles, it, too, must here be discussed. This effect, which, according to Musatti, was discovered by Benussi in 1921, will be described later. Musatti refers to both these effects as 'stereo-kinetic' phenomena. They bear resemblance to the kinetic depth-effect, whose properties have more recently been investigated, and the question arises whether they can be considered as instances of this effect.[5]

In brief, the kinetic depth-effect (*KDE*) can be obtained when the shadow of a solid object in rotation is observed. Under these circumstances usually a solid form is perceived similar to that which casts the shadow. This happens even if that form is so chosen that, when it is presented stationary, its shadow is seen as a plane figure. The shadow of the turning

FIG. 6. AN ELLIPSE ON TURNTABLE

---

[5] Hans Wallach and D. N. O'Connell, The kinetic depth effect, *J. exp. Psychol.*, 45, 1953, 205-217.

object changes its shape continuously and this deformation is the cause for the three-dimensional appearance of the shadow. A rigid three-dimensional form is seen instead of the continuously distorted plane figure that is actually given. Only a turning motion of the perceived rigid form reflects the given distortions. This motion usually will closely resemble the rotation of the object that casts the shadow, although its direction may be reversed. The KDE can also be obtained with the shadow of a plane figure that turns in space. The effect here is to cause a rigid plane figure, turning in depth, to appear in place of the distorting shadow.

There is, of course, an important difference between the 'stereo-kinetic' phenomena and the *KDE;* namely, that in the case of the former no deforming patterns are objectively given. Rather, the deformations that are the antecedent of a depth-effect are perceived merely as a result of another perceptual effect which causes the given rotating patterns to distort. Yet, it is possible that the transformation from a percept consisting of a deforming figure located in the plane of the turntable into one consisting of a form arranged in depth is an instance of a *KDE.*

To test the possibility that the stereo-kinetic phenomena are instances of the KDE we have to ask whether they conform to the characteristics of this effect. If stimulus conditions are adequate, the KDE will occur without suggestion and irrespective of set. This criterion emerged in previous research on the KDE.[6] It was found that a three-dimensional effect will occur regularly and spontaneously when the shadow or the retinal image of a turning object shows contours or lines that change simultaneously in direction and length. On the other hand, forms whose deformations consist merely in a lengthening and shortening of contours will not regularly produce depth-effects. For instance, the shadow of a T-shaped wire figure that turns back and forth about its vertical axis will be seen turning as a rigid form only by one out of three naïve Os; the others will see the horizontal line expand and contract in the plane of the screen. An appropriate suggestion or prior presentation of a pattern that does produce a KDE, however, will cause almost all Os to see a turning T-figure.

The following experiment is an attempt to test in similar fashion whether Musatti's effect occurs regularly and spontaneously.

Procedure. A black cardboard disk 20 in. in diameter on which a white ellipse was pasted (Fig. 6) was turned about its center at a rate of about 20 r.p.m. The turntable stood on the floor and the disk revolved in a horizontal plane about 6 in. above the floor. O was directed to stand next to the turntable and to look straight down on it. The illumination was indirect and of moderate brightness, a condition

[6] Wallach and O'Connell, *op. cit.,* 209-211.

favorable to the emergence of the effects to be observed. Forty-seven Os participated. They were students and some staff members of Swarthmore College. None was familiar with the effects to be observed.

O began his observation with both eyes open and was told to describe the way the white figure looks as it moves around the disk—and to report any changes in its appearance as they occur. After 30 sec. O was asked to cover one eye with his hand and to continue to describe the way the figure looked to him. When O failed to describe the ellipse as a disk rising out of the plane of the turntable at a tilt, E asked the following question: "Can the figure be seen as a disk rolling around on its edge?"

The Os failing to observe the Musatti effect spontaneously were divided into three groups. One group (17 Os) was asked the question after 10 sec. of monocular observation; a second group (17 Os) was asked the question after 30 sec.; and the third group (13 Os), after 60 sec. After the question was asked the observation-time was extended 45 sec.

Results. Sooner or later during the binocular observation period all 47 Os saw the ellipse deforming, but none of them saw it as a disk rising out of the plane of the turntable. Inasmuch as a KDE can be obtained regularly with binocular observation of a deforming shadow, this in itself is an important result; it differentiates between Musatti's effect and the KDE. During monocular observation only 6 Os reported seeing the tilting disk prior to the suggestion and 7 Os never saw it. This leaves 34 Os who saw the tilting disk only after the suggestion had been given. Of the 30 Os for whom the suggestion was to be delayed until 30 or 60 sec. of monocular observation had passed, 5 saw the tilting disk prior to the suggestion, 5 never saw it and 20 saw it only after the suggestion had been given. The fact that 17 of these 20 Os reported seeing the tilted disk within 5 sec. after the suggestion had been made further argues that these reports are connected with the suggestion.

Some reader may wish to question whether Os who reported a tilting disk only after the suggestion was given actually perceived it. The E who received these reports believed that O described a real perceptual experience. The descriptions usually included features that were not given or were not logically implied in the suggestion. After O had reported a tilting disk he was asked to move back 2 ft. from the turntable and look obliquely down on it. When the turning pattern is viewed from such an angle, the disk may change its slant as it seems to roll around the turntable, and such an experience is in agreement with conditions of stimulation.[7]

---

[7] Assume a real disk is rolling about the turntable leaning toward its center. (All our Os saw the disk tilted in this direction.) If the disk were maintaining a constant angle of slant, say of 45°, it would intercept the line of regard more nearly at a right angle when it is in the near position where its slant is away from O than when it is in the far position where it leans toward O. The shape of its elliptic retinal

O was asked, "Does the disk appear to stand up more vertically when it is nearest to you or when it is opposite you?" Of the 34 *O*s who saw the tilting disk only after the suggestion had been given, two did not notice changes in slant, 30 reported those changes that should be seen if the perceived slant fitted conditions of stimulation and two *O*s reported the opposite changes. Thus the great majority of *O*s noticed features of the perceptual process that were not contained in the suggestion or deducible from it without an intimate knowledge of the complex geometrical situation; they could only be 'perceived.' (Failure to report the correct changes does not imply that *O* did not actually see the tilted disk; it merely means that the oblique direction of viewing the turntable failed to have its effect on the slant of the disk.)

Thus it sems safe to conclude that the majority of naïve *O*s see the tilted disk only when an appropriate suggestion is given. By no means can it be claimed that Musatti's effect occurs regularly and spontaneously. It is in this respect strikingly different from the *KDE*.

BENUSSI'S EFFECT

Benussi's effect can easily be obtained when a pattern like Fig. 7 is turned about its center and is observed monocularly. Quite soon the rings will appear to be arranged in depth. Usually they seem to form a cone rising out of the plane of the turntable. The top of the cone performs a slight circular motion, but apart from this a rigid solid form is seen.

Here, too, the three-dimensional stage is preceded by another illusory perception: The ring pattern, still in one plane, fails to rotate as a whole. Instead the pattern is in continuous deformation. Although none of the rings appears to turn, the inner ones circle about in relation to the outer ring and the extent of this motion is different for each, with the innermost ring moving most. This peculiar experience again finds its explanation in the tendency of a rotating circle to appear to stand still, which was discussed above. If rotation is eliminated from their motion, only the slight revolution of the inner rings will be visible, which is the result of their eccentric location on the turntable. Since their eccentric locations differ in degree and therefore their revolving motions differ in extent, they perform this motion independently of each other, and the pattern appears to undergo distortion.

In the three-dimensional stage an almost rigid solid, a truncated cone,

---

projection would therefore change; the retinal image would be more circular when the disk is in the near position. Actually, however, the retinal image hardly changes, because an ellipse in the plane of the turntable is objectively given. Such an unchanging retinal image can be produced by a real disk only when its slant is steeper in the near position and more nearly horizontal in the far position.

replaces the deforming plane pattern. Again the question arises whether this is the result of a *KDE*. An experiment was performed to test whether Benussi's effect occurs regularly and spontaneously. Since, however, Fig. 7 is sometimes seen in three dimensions when given at rest, a different pattern of circles was designed which, when stationary, remains in a plane for most *O*s and this pattern was used in our experiment.
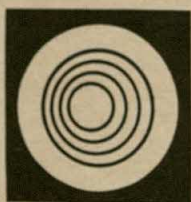
FIG. 7. FOUR RINGS
ON TURNTABLE
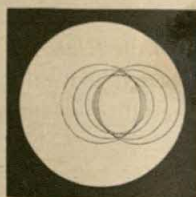
FIG. 8. SIX OVER-
LAPPING RINGS ON
TURNTABLE

Procedure. The pattern of Fig. 8 was fastened to the turntable and presented under three different conditions: *O* first described it at rest; then it was rotated at 20 r.p.m. and observed binocularly, *O* being asked to describe what he saw and report any changes that might occur; then, if he failed to report a three-dimensional form within 30 sec., he was asked to cover one eye. After 30 sec. of monocular observation he was asked whether the figure could be seen as three-dimensional. Twelve naïve *O*s participated.

Results. All *O*s described the stationary pattern as a plane figure. During monocular observation of the rotating pattern 11 *O*s reported the three-dimensional form, which in the case of Fig. 8 resembled a bedspring, spontaneously and only one needed the suggestion. Even more strikingly 10 of the 11 *O*s saw it during binocular observation. Thus the result for the pattern that produces Benussi's effect is strikingly different from our results for the ellipse. Benussi's effect was almost always spontaneously obtained whereas, in the case of the ellipse, of a comparable group of 25 *O*s who ultimately saw the tilting disk, 20 needed a suggestion to do so. Moreover, 10 out of 12 *O*s obtained Benussi's effect with binocular observation, whereas the tilting disk was never seen under this condition.

Our results show that Benussi's effect occurs regularly and spontaneously. It resembles in this respect the *KDE*. The question arises whether other criteria of the *KDE* fit Benussi's effect. It has been mentioned that the *KDE* occurs where contours change simultaneously in direction and length. It should be added that this also applies to intervals between objects or points: if an imaginary line connecting them undergoes these changes

simultaneously they will appear to move at different depth.[8] To be sure, the rings and circles of the Benussi patterns do not present the eye with distinct parts. Nevertheless, such parts are perceived. In spite of the rotation of the whole pattern each circle is clearly seen to be at rest, that is, to maintain its orientation, again as the result of an imparting of identity. In the present case, this effect accounts for seeing a part of one circle move in relation to a part of another circle. Imaginary lines connecting such parts change in length and direction (Fig. 9) and thus the deformations meet the conditions of the *KDE*.

It has been stated above that in the case of the ellipse, too, the deformations are the result of an imparting of identity. Here, however, the result
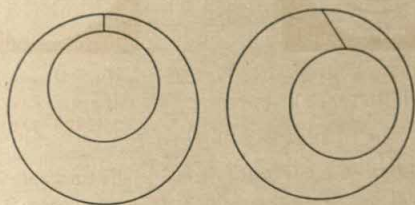
FIG. 9. TWO CIRCLES BEFORE AND AFTER CLOCKWISE ROTATION OF 90°
The line that connects the top points of the circles changes both in length and direction.

is usually incomplete; the figure, though deforming, is also seen to turn as it revolves about the center of the turntable. Whether this is the reason for the failure of the depth effect to occur spontaneously in the case of the ellipse remains to be seen.

SUMMARY

A rotating circle seems to stand still, a rotating spiral seems to expand or contract, and an ellipse displays flowing deformations when it is turned. All these observations were found to be related. They were explained as the result of subjectively identical points along the contours of these figures that arise in relation to the absolute directions of visual space.

When such illusory deformations are observed monocularly, a change may take place and so-called stereo-kinetic phenomena may be observed. Instead of the deforming patterns, a solid form or a rigid plane figure moving in depth may be seen. These phenomena were compared with the kinetic depth-effect. It was found that one of these phenomena (Benussi's effect) can indeed be regarded as a case of the kinetic depth-effect, whereas one obtained with an ellipse is of a more complex nature.

---

[8] Wallach and O'Connell, *op. cit.,* 214.

SPEED OF PERCEPTION AS A FUNCTION OF
MODE OF REPRESENTATION

By T. A. RYAN and CAROL B. SCHWARTZ, Cornell University

In designing illustrations for a textbook, a training manual, or a film strip, the author has the choice of several different modes of representation, the four principal modes being photographs of the object, shaded drawings, line drawings, or cartoons. Differences in effectiveness among these methods could be expected to be most marked where critical characteristics of the objects are three-dimensional. The study to be reported is a preliminary research into the comparative effectiveness of these methods. The research is preliminary and the conclusions are tentative because we could study only a limited number of objects and their representations, and because it was necessary to develop a methodology for the purpose.

We have assumed that an important criterion for the effectiveness of an illustration is the speed with which the relevant details can be perceived. That is, we decided to measure effectiveness in terms of the tachistoscopic threshold for specific aspects of the illustration. We believe that there is good practical justification for making speed of perception our criterion. If a lecturer is using a film strip to illustrate his presentation, the sooner his audience can see the relationships he is talking about, the sooner are his hearers ready to understand his explanations. Similarly, a textbook which contains quickly perceived illustrations will be read faster. An illustration which requires a long perception-time must contain confusing or conflicting elements and is therefore less directly relevant to its purpose.

In many of the experiments on the perception of form in which the tachistoscope is employed, the threshold which is obtained is the threshold of recognition or identification of the whole object. In such studies, the interest has centered upon the silhouette or outline of the object—its two-dimensional form.[1] In most cases

[1] Tachistoscopic studies relevant to the present study may be classified into three main groups. In one category are those concerned with the temporal 'stages' of perception. One of the more recent examples is: A. G. Douglas, A tachistoscopic study of the order of emergence in the process of perception, *Psychol. Monogr.*, 61, 1947 (No. 287), 1-133. Here the three-dimensional aspects of the objects presented, if such aspects are present at all, are considered as merely another set of details. Secondly, there is a group of studies specifically related to three-dimensional perception. These have been directed primarily at evaluating the effectiveness of such factors as stereoscopy in the absence of eye-movements. In such experiments, the tachistoscope is used merely as a method of controlling possible eye-movements. See,

where illustrations are used to aid in exposition of technical, scientific, or mechanical material, however, the perceiver already knows what the object's identity is. He has already been told that the object he will see is, for example, a lathe, a diagram of the nervous system, or an illustration of a method of tying knots. What he looks for in the picture is not the identity of the object, but its detailed structure. It was this latter aspect of perceiving that we wished to study, so we measured time-thresholds for perceiving specified characteristics of the objects depicted, with S already informed as to the identity of the object as a whole.

Procedure: Materials. Three objects, each in four different positions, were chosen for this study. One experimental series with a distinct group of Ss was devoted to each of the objects. The objects were (1) a human hand, against a light background, in four different orientations; (2) a group of five electrical knife switches (double throw) with one switch open and the rest closed in each illustration, again with four different combinations of the positions of the switch handles; (3) a cut-away model of the valves of a steam engine at four different stages of the cycle. Each of these 12 combinations was reproduced in 4 different ways, making a total of 48 illustrations. Figs. 1, 2, and 3 illustrate the four representations of Position 1 for each object.

(1) *Photographs.* Each object, in each position, was first photographed by a professional photographer under E's supervision. It was specified that all photographs were to be taken with the objects against a light background, using the best possible front and side lighting, but without back lighting. Every precaution was taken to standardize the degree of contrast and the density of each print. The other forms of illustration were then drawn from these large prints.[2]

(2) *Shaded drawings.* These were literal reproductions from the photograph except that the number of gray tones was reduced. The gray tones were controlled by making the drawings upon Craft Tint paper No. 277 which makes it possible to produce two controlled shades of gray by uniform stippling. All lines were kept at a uniform width.

(3) *Line drawings.* These were tracings of the essential outlines of the photograph. Uniform line-width was also maintained here.

(4) *Cartoons.* Using the 'animation' technique, we distorted the figure to emphasize the essential spatial relationships involved. We also aimed at producing these distortions without modifying the total silhouette of the figure, because we did not wish the discrimination to be based upon the silhouette. Some slight modifications of the silhouette were necessary, but they were kept at a minimum.

Slides. The 48 illustrations were then prepared in the form of 2 × 2 in. slides, again controlling the contrast and density for illustrations of a particular type, and keeping the size of the images constant for any given object.

Method of presentation. The illustrations were projected for controlled time-intervals upon a screen 68 in. from *S*. The size of the image was 22 × 28 in. Projection was controlled by means of shutters before the lenses of the two Kodak projectors. The shutters were operated by solenoids controlled by a Stoelting Interval Timer. One projector furnished a preëxposure field (a blank slide with a fixation-point). This field was occluded simultaneously with the presentation of the stimulus-illustration.

The time-control for the complete apparatus was calibrated by means of a photo-electric cell with a sensitive relay operating a Standard Electric Time Clock. The reliability was very high, giving variations of less than 0.01 sec. for any given setting of the timer, and a standard deviation of about 0.002 sec. Because of possible lags in the system, the absolute times which are given in our results may be in error by small constant amounts, but the stability of the calibration shows that the variable error is inconsequential.

The lower limit of the time-control was 0.02 sec., and preliminary trials showed that this was too slow for many *S*s with the bright image which was given by the projectors at this distance. The brightness of the image was therefore reduced by pairs of polaroid filters over each projector lens. Brightnesses were measured with no slide in the projector and were set at 0.18 ft. lamberts for the preëxposure-field and 0.13 ft. lamberts for the exposure-field. With indefinite time of exposure, there was no difficulty in seeing the details of the pictures at this brightness.

For each slide, the presentation-times were given in a single ascending series which continued until *S* gave the correct response. Each series began with 0.02 sec. To keep the number of presentations for a single figure to a reasonable limit, the series consisted of steps of 0.01 sec. up to 0.10 sec.; and then steps of 0.10 sec. up to 1 sec. In the few cases where a correct response was still not forthcoming, the remaining steps were increments of 1 sec. This procedure was necessary because of the use of only a single series for each picture. Where the threshold was less than 0.1 sec. small steps were necessary. To continue with these small steps all the way up to 3 or 4 sec. would have meant a prohibitive number of presentations of the object, and we do not know what the effect of so many repetitions might be.

The ideal design would be only one presentation for a given illustration for a given *S*, with the time selected at random. This procedure would require a prohibitively large number of *S*s, especially since the experiment was meant to be only exploratory.

In presenting the results we shall speak of our recorded results as 'thresholds' even though each is based upon but one ascending series. It is the only threshold which can be obtained, since any subsequent presentations of the same picture would be influenced by knowledge of that picture. There is, of course, a considerable element of chance in each of these thresholds but we can allow for this by means of the statistical analysis. The method of spacing the times in the ascending series also had the effect of skewing the distributions of the thresholds, an effect which also had to be taken into account in treating the results.

Each series utilized the 16 illustrations of one object—4 positions of that object and 4 modes of illustrating each of these positions. The 16 slides were presented in a different random order for each *S*, with the random orders selected in such a way that each slide appeared once in each position of the series for one of the *S*s.

Thus, for example, the cartoon of the hand in Position 1 was presented first in the series for one S, second in the series for another S and so on for each of the four positions. Thus any learning effect was equated for all of the illustrations.

The procedure in each of the series was identical, except that in the first two series the observations were repeated in reversed order. This reversed presentation was found to contribute little additional information and was not included in the third series.

Method of report. After each presentation of each slide, S was asked to reproduce the position which he had seen. For the hands, he was to place his own hand in the position he saw. For the switches he attempted to name the particular switch which was open, and for the steam valves, to state which stage of the cycle was shown. In the latter case, before each S began his observations he was shown the model itself and learned designations for the different positions. During the series, in any case where S had difficulty with the verbal identification of the position, he was permitted to demonstrate on the model the position he had seen. Thus S knew in advance the kind of object which was to be shown, and his only task was to discriminate the position in which that object was depicted. In the case of the switches and valves there was no ambiguity in determining when S was giving the correct response. There was a limited number of choices and no intermediates were used. In the case of the hand, the reproduction was scored correct if the general position of the fingers was correctly reproduced. If some of the fingers were in correct position while others were clearly out of position, the reproduction was counted as wrong.

Results. Tables I, II, III, and IV give the average thresholds for 16 Ss for each of the illustrations and for all objects taken together. The bottom

TABLE I

MEAN THRESHOLDS (IN SEC.) AND MEAN RANKS

(Hands)

Position	Mode of reproduction			
	Photo	Shaded drawing	Line drawing	Cartoon
1	.038	.046	.128	.051
2	.033	.032	.052	.049
3	.050	.046	.054	.047
4	.033	.039	.048	.039
Mean	.039	.044	.070	.046
Mean rank	1.75	1.88	3.91	2.47

line of each table gives the mean rank-order of the different modes of presentation. This mean rank was obtained by finding the four average thresholds for each S, ranking these for each S, and then finding the mean of all of the ranks assigned to each mode of presentation. The reason for this procedure will become evident when we discuss the significance tests for these results.

A preliminary analysis of order or practice effects showed these to be negligible. The data were therefore grouped for a 3-variable analysis of variance (Ss, poses, and methods of presentation) but the distributions

TABLE II

Mean Thresholds (in Sec.) and Mean Ranks
(Switches)

Position	Mode of reproduction			
	Photo	Shaded drawing	Line drawing	Cartoon
1	1.133	.304	2.564	.680
2	.050	.856	.146	.084
3	.459	.519	.975	.301
4	1.119	1.198	.989	.086
Mean	0.690	0.719	1.169	0.288
Mean rank	2.56	2.53	3.41	1.5

TABLE III

Mean Thresholds (in Sec.) and Mean Ranks
(Valves)

Position	Mode of reproduction			
	Photo	Shaded drawing	Line drawing	Cartoon
1	.148	.248	.623	.120
2	.403	.131	.511	.119
3	.224	.628	.170	.126
4	.165	.258	.198	.683
Mean	.235	.316	.375	.262
Mean rank	2.719	2.625	2.969	1.688

TABLE IV

Mean Thresholds (in Sec.) and Mean Ranks
(All objects)

Mean	Mode of reproduction			
	Photo	Shaded drawing	Line drawing	Cartoon
Time	.32	.36	.54	.20
Rank	2.34	2.32	3.45	1.89

were found to be markedly skewed. Also it was found that non-parametric analysis gave a clearer picture of the results than did ordinary analysis of variance. The pattern of ranks was tested for significance by methods developed by Friedman[3] and the results of these tests are reported in Table V.

---

[3] Milton Friedman, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *J. Amer. Statistical Ass'n.,* 32, 1937, 675-701.

As the tables show, the over-all tendency is for the line drawings to be the most difficult to perceive accurately, while the cartoons are most quickly perceived. Photographs and shaded drawings are about equivalent and fall somewhere between the cartoon and the line drawing. In spite of individual

TABLE V
SIGNIFICANCE OF DIFFERENCES AMONG REPRESENTATIONS

Object	Rank tests		
	χ_r^2	df	P
Hands	28.13	3	< .001
Switches	17.53	3	< .001
Valves	9.06	3	< .05
Mean all objects	38.36	3	< .001

differences and variations from object to object, the over-all pattern is significant.

To illustrate the variations from object to object, and from pose to pose, Table VI shows the rankings of the mean thresholds within each pose. The line drawings had the highest mean threshold for 8 out of the 12 object-

TABLE VI
RANKS OF MEAN THRESHOLDS FOR EACH POSE

Object	Pose	Photo	Shaded drawing	Line drawing	Cartoon
Hands	1	1	2	4	3
	2	2	1	4	3
	3	3	1	4	2
	4	1	2.5	4	2.5
Switches	1	3	1	4	2
	2	1	4	3	2
	3	2	3	4	1
	4	3	4	2	1
Valves	1	2	3	4	1
	2	3	2	4	1
	3	2	4	3	1
	4	1	3	2	4
Rank (frequency)	1	4	3	0	5
	2	4	2.5	2	3.5
	3	4	3.5	2	2.5
	4	0	3	8	1

pose combinations, and the line drawings never had the lowest thresholds for any pose.

The interaction between object and pose was significant at the 1% level for the hands and the 5% level for the valves, as tested by ordinary analysis

of variance. It was not possible to test these interactions by means of non-parametric methods, but the indications are that there is a substantial inter-action to be reckoned with—that the value of a particular mode of repro-duction will vary from object to object.

Thus there is a possibility that the over-all results could be due to the
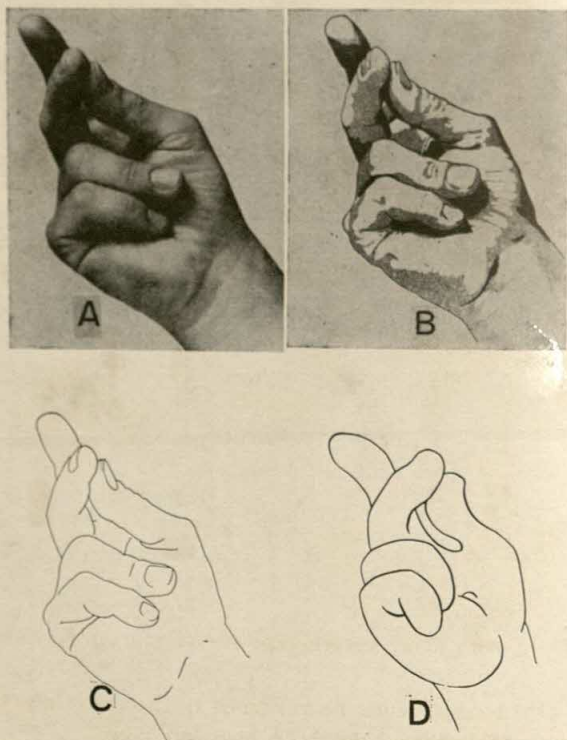
FIG. 1. FOUR REPRESENTATIONS OF THE HANDS IN POSITION 1
(A) Photograph; (B) Shaded Drawing; (C) Line Drawing; (D) Cartoon

selection of objects and poses to be depicted, but we do not believe this to be the case. Our objects and poses were not chosen as especially suitable for any particular method of reproduction, but rather as objects which are of the kind which might appear in training manuals, and which were readily available. Of course, they do not constitute a truly random sample of objects of the kind we were interested in (see specifications in introduction), but neither can we see how their selection could have been biassed, especially as

we ourselves had no predictions as to the outcome. If, on this basis we treat the 12 object-pose combinations as though they were a random sample of such combinations, the fact that line drawings had the longest thresholds
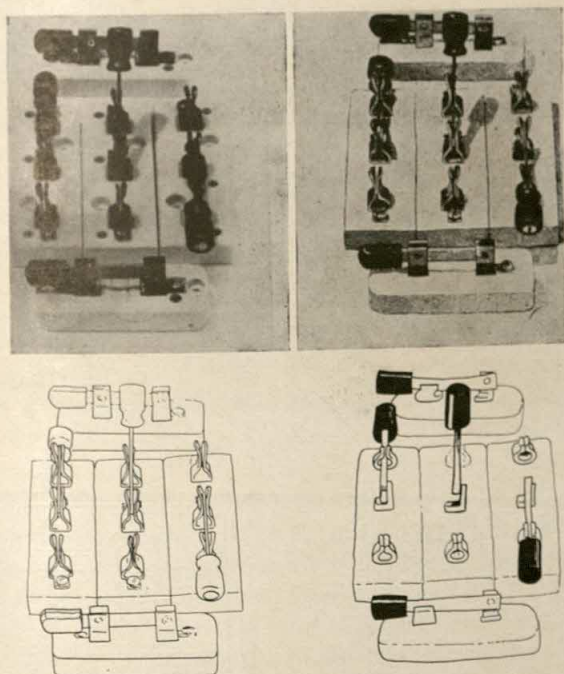
FIG. 2. FOUR REPRESENTATIONS OF THE SWITCHES IN POSITION 1

in 8 out of 12 cases cannot be attributed to chance. (Only 3 would be expected to have the longest thresholds by chance alone.)

Thus we believe that it is safe to conclude that our results apply to objects of the kind we selected, where the discrimination is of the relative position of a complex object in three dimensions, and where the silhouette is relatively ineffective as a clue to the position. For such objects, line drawings are the least effective mode of presentation, and cartoons the most effective.

The exceptions to this rule seem to be difficult to account for, except on the basis of chance variations. For example, Positions 3 and 4 for the valves differ mainly in being right-left reversals of the same position. Yet, in one case the cartoon has the lowest threshold, and in the other, the photograph is the lowest.

The fact that *S* knew in advance what the object was going to be must also be stressed as an important condition for these results. It is probably the reason why there was no observable practice effect running through the series of observations. It also means that the thresholds were shorter than would ordinarily be found when complex objects like these are presented without preparation. We believed it to be a desirable condition for the experiment, since we wished to compare different forms and illustrations of the same object. We wished to do this with the same *Ss*, in order to be
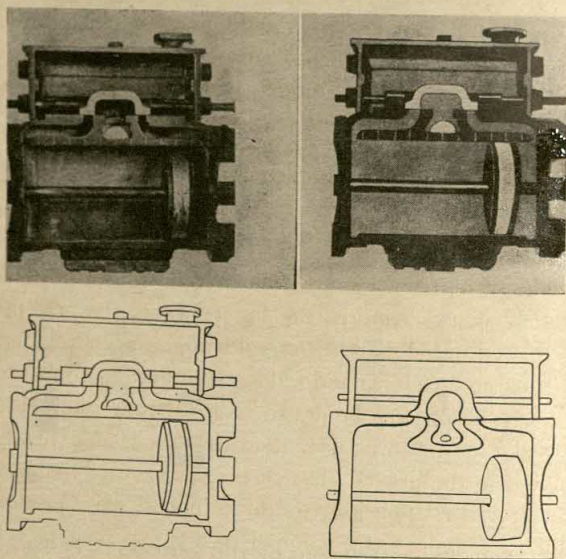
FIG. 3. FOUR REPRESENTATIONS OF THE STEAM VALVES IN POSITION 1

able to control the large individual differences which always increase the sampling error in such experiments. If *S* had not been prepared in advance, the first several trials would have been entirely different from the rest because he would eventually 'catch on' to the fact that all of the illustrations involved the same object.

It may be that the differences between modes of illustration would be even larger if all of the thresholds involved objects which came to *S* without advance information. Nevertheless, we felt it desirable to place this restriction upon our results because we wished to separate the problem of identifying the object, from detailed perception of its position.

Repetitions of series. The first two series of slides (the hands and the

switches) were presented twice to each S, with the second presentation being in reverse order. The purpose of the repetition was to explore the possibility that certain methods of reproduction would be more easily perceived than others, when seen a second time. In other words, *improvement* in speed of perception could be taken as another criterion of adequacy of mode of illustration.

As a result of these repetitions we had, therefore, a difference-score for each slide and for each S. We could find no pattern relating these difference-scores to mode of illustration, however, in either of the first two series. Consequently, as time was pressing, the duplicate presentation was omitted in the last series (valves). Another experimental design would probably be needed for satisfactory study of improvements in speed of perception as a function of previous exposure to the same stimulus.

CONCLUSIONS

Although this study was meant to be only a preliminary investigation, significant differences in thresholds for time of perception were found for different modes of representation. On the average, for the objects and poses we used, line drawings required the longest times for perception, while cartoons were seen in the shortest times. Photographs and shaded drawings were about equal to each other and fell between line drawings and cartoons. While these conclusions can be drawn with certainty only with reference to the particular objects we used, there is considerable justification for extending them to the general class of objects we used—complex objects in which the crucial relationships are three-dimensional. Certainly we have enough evidence to point to the value of further experimental series of this kind which would extend the sample of objects beyond that of this study.

As a preliminary study, this experiment also served the purpose of evaluating a method for the study of this kind of problem. By including in a single experimental series only depictions of the same object in various poses and in varied modes of presentation, it was possible to obtain thresholds for 16 different slides for each S. Practice effects were found to be negligible, with the result that we could gain the statistical advantages of intra-S comparisons, thus increasing the sensitivity of the experimental design and reducing the costs of the series.

THE INFLUENCE OF VISUAL STIMULATION ON KINESTHETIC FIGURAL AFTER-EFFECTS

By Robert Jaffe, Mount Sinai Hospital, New York

Ample evidence exists to show that the perceived size, shape, and spatial position of an object may be influenced by preceding stimuli.[1] One alteration in the subjective dimensions has been termed "figural after-effects" and has been correlated with the characteristics of the antecedent stimulation. In all studies to date, this prior stimulus, as well as the stimuli used to test for presence of the after-effect, has been restricted to presentations within one modality, *i.e.* visual stimulation has been used to elicit visual after-effects and kinesthetic stimulation to elicit kinesthetic after-effects. As yet, no study has attempted to demonstrate after-effects across modalities; for example, from vision to kinesthesis or vice versa.

The significance of a possible intermodal effect on perceived size could be considered in terms of Köhler's theory of satiation.[2] Three basic assumptions may be found in this theory. The first is that direct-current potentials are an essential physiological concomitant of object-perception.[3] Secondly, prolonged stimulation supposedly results in areas of altered electrotonus in the cerebral cortex, due to the differential polarization of cerebral tissue by direct currents, *i.e.* satiation. Thirdly, Köhler assumes that these areas of altered electrotonus are restricted to the primary projection system, and, at least for vision, are topologically related to the form as well as the spatial position of the perceived figure, *i.e.* isomorphic representation. Inspection (or palpation) of a stimulus-figure sets up localized electrotonic fields that produce changes in cortical conductivity.

[1] Wolfgang Köhler and Dorothy Dinnerstein, Figural after-effects in kinesthesis, *Misc. Psychologica Albert Michotte*, 1947, 196-220; J. J. Gibson, Adaptation, after-effect, and contrast in the perception of curved lines, *J. exp. Psychol.*, 16, 1933, 1-31; Köhler and Hans Wallach, Figural after-effects: an investigation of visual processes, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

[2] Köhler and Wallach, *op. cit.*, 315-357.

[3] Köhler, Relational determinants in perception, in L. A. Jeffress (ed.), *Cerebral Mechanisms in Behavior (The Hixon Symposium)*, 1951, 200-243. There is some direct electrophysiological evidence to support this assumption. See Köhler, R. Held, and D. N. O'Connell, An investigation of cortical currents, *Proc. Amer. Phil. Soc.*, 96, 1952, 290-330.

As a result of this altered conductivity, the distribution of the potentials set up by subsequent stimuli is distorted. The behavioral consequences of this entire process are the reported distortions in perception.

Questions might be raised in regard to Köhler's assumption of the isomorphic representation of electrotonic activity in the primary projection-areas of the cortex. If satiation is restricted to the primary projection-area, figural after-effects should only be demonstrable when both the antecedent stimulus and the stimuli used to test for the after-effect are presented to the same modality. The present study was designed to investigate this third assumption. In behavioral terms, the hypothesis to be tested was that the figural after-effect is a phenomenon specific to any given modality.

EXPERIMENT I

Subjects. The 20 Ss in this study were drawn from laboratory personnel, residents in neurology and psychiatry, graduate students in psychology, medical students, and laboratory assistants; none had prior experience with testing conditions used to elicit figural after-effects. Half of the Ss served as an experimental group, the remaining 10 as controls.

Materials. The apparatus used to elicit kinesthetic figural after-effects was the modification of the Köhler-Dinnerstein apparatus used in our earlier study.[4] A strip of aluminum 2 in. wide was used both as standard and as interpolated stimulus in this investigation; *i.e.* the same strip was used throughout all phases of the kinesthetic testing procedure. A strip of aluminum graduated in width from 1 to 3 in. was the comparison scale. Visual stimuli consisted of three strips of heavy white paper which could be pinned to a black felt screen. The strips were 24 in. long and 1, 2, and 4 in. wide.

Procedure. The control group was tested under conditions of tactile-kinesthetic stimulation alone. S stood, blindfolded, between two tables. One table supported the 2-in. standard strip, the other the graduated scale. The test consisted of three steps: (1) standardization; (2) interpolated stimulation; and (3) a test-judgment of width. S performed the standardizing judgment by grasping the 2-in. strip between the thumb and fingertips of the left hand, while similarly grasping the scale with the right.[5] E moved the scale until S reported the width of the two strips to be equal. Six readings, three ascending and three descending, were taken. S performed Step 2, interpolated stimulation, by running the hand previously used for holding the standard up and down its length for 1 min. Following this, the judgment of width was again performed (Step 3). The entire procedure was repeated after a 5-min. rest-period.

The experimental group followed the same procedure as the control group with the following exception. Instead of being blindfolded, S wore goggles that re-

---

[4] Robert Jaffe, Kinesthetic after-effects following cerebral lesions, this JOURNAL, 67, 1954, 668-676.

[5] Other data obtained with a similar group of normal Ss had shown that handedness did not influence performance in the test of kinesthetic after-effects. For this reason no control for hand-preference was made in this study.

stricted his field of vision to the surface of a 60 × 60-in. black screen 6 ft. distant. During the standardization, S was instructed to gaze at a strip of white paper, 2 in. wide, that was pinned to this screen with its long axis vertical. During the two periods of interpolated stimulation, this 2-in. strip of white paper was replaced by a white strip either 1 in. or 4 in. wide, and S was instructed to look at the white strip while moving his left hand along the 2-in. aluminum standard. One half of the Ss regarded the 1-in. strip during the first period of interpolated stimulation and the 4-in. strip during the second. This order was reversed for the remaining Ss. In Step 3, the test judgment, the Ss were retested for the apparent width of the tactile standard (still 2 in. wide) while looking at the same 2-in. white strip of paper which had been presented during the initial standardization.

Results. The results obtained with both groups are shown in Fig. 1. The points represent the group mean values for the change in the tactually
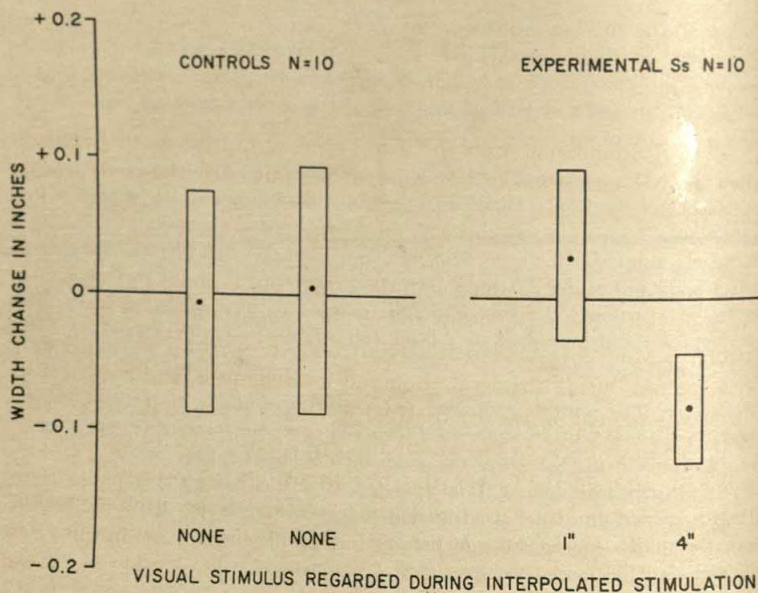
FIG. 1. CHANGE IN APPARENT WIDTH OF KINESTHETIC STANDARD
Points with the bars represent the mean change of 20 Ss in the apparent width of the kinesthetic standard after interpolated stimulation; the rectangular bars represent the range of one standard deviation each side of the mean.

perceived width of the standard object; the bars represent a range including one standard deviation on either side of the mean. The results for the control group show that there was no major change in apparent width of the 2-in. standard after interpolated stimulation with the same strip. The inter-individual variability of this group is large. By contrast, the Ss of the experimental group, who *had* experienced concurrent visual stimulation

during the tactile-kinesthetic testing procedure, did show notable changes in the apparent width of the standard. The standard appeared to increase in width after visual stimulation with the narrow strip, and to decrease in width after visual stimulation with the wide strip.

When evaluated by a *t*-test (see Table I), there was a significant underestimation of the standard after visual stimulation with the 4-in. white strip.

TABLE I

Tests for Significance of After-effects within Groups

Ss	Mean Distortion	SD	t	P
Controls (N = 10)				
Test 1	− .01	.08	.37	.80
Test 2	.00	.09	.07	.90
Experimental (N = 10)				
Test 1 (1-in. strip)	.03	.06	1.44	.20
Test 2 (4-in. strip)	− .08	.04	5.65	< .01

After visual stimulation with the 1-in. white strip, the mean judgment shows an over-estimation of the standard but this effect does not achieve significance.[6]

Experiment II

The positive results obtained with the experimental group in Experiment I suggested that the relationship among the visual stimuli was the crucial variable in producing the kinesthetic after-effect. Since a 2-in. white strip was inspected during standardization and test-judgment, and a 1- or 4-in. strip was presented during interpolated stimulation, one could hypothesize that the contrast between the size of the strips was the factor inducing the kinesthetic after-effect. To test this hypothesis, the experimental Ss were retested (after an interval of 3-4 mo.) with the 1- or 4-in. strip again fixated visually during the interpolated period of stimulation. In this retest situation, however, *no* visual figure was presented during the initial and final width-judgments. Under these conditions no significant change in the apparent width of the kinesthetic standard object could be demonstrated (see Table II).

The results of both experiments can be summarized by saying that concurrent visual stimulation can induce a significant kinesthetic figural after-effect, but that this after-effect appears only when contrasting size relationships exist between the visual stimuli.

[6] In my previous study it was noted that the 1-in. interpolated *kinesthetic* stimulus-object elicited a significantly smaller after-effect than the 4-in. object. See Jaffe, *op. cit.*, 617.

DISCUSSION

The results of this study are contrary to certain predictions that were made on the basis of satiation theory. Köhler and Wallach originally postulated that satiation takes place in the primary projection-area of the cerebrum.[7] According to this point of view, one would not predict that stimulation in one modality would result in distorted object-perception in another. If the satiation theory is to be in-

TABLE II

AFTER-EFFECTS SHOWN BY Ss IN EXPERIMENT II

Tests	Mean Distortion	SD	t	*	P
Test 1 (after 1-in. strip)	−.02	.06	1.05		.40
Test 2 (after 4-in. strip)	−.01	.03	.55		.60

voked to explain our results, the theory would have to be extended and modified, perhaps in line with Köhler's remark that figure currents might spread across wide areas of the brain, and thus account for remote effects.[8] Such modifications would require fundamental changes in the physiological postulates of satiation theory, however, inasmuch as isomorphic projection is one of those postulates.

Köhler's assumption that the neural events take place in the primary projection-areas is also difficult to reconcile with some of the data obtained in our previous study.[9] One group of Ss in that study had brain injuries which produced definite neurological evidence of somato-sensory defect and therefore could be assumed to have received damage to the primary projection-system. In spite of this damage, the figural after-effects in these Ss were indistinguishable from those of normal controls.

The theory of satiation was first elaborated in regard to two-dimensional after-effects in vision, and within this framework it showed a high predictive capacity. Köhler and Dinnerstein admit, however, that the earlier physiological model is not adequate to explain after-effects in kinesthesis and in the third dimension of visual space.[10] Köhler indicates that difficulty arises in trying to account for perception in three-dimensional space in terms of the concept of a two-dimensional cortical substrate. For the same reasons, Köhler's reservations may apply to more recent studies which demonstrated after-effects in auditory and kinesthetic space.[11] The results of the present and my previous studies add to the body of the data that cannot be accounted for in terms of the original formulation of the satiation theory-

---

[7] Osgood and Heyer offer another detailed theory of the physiological basis of figural after-effects. Their theory, however, invokes an even more specific mechanism of point-to-point projection than that of Köhler and Wallach. See C. E. Osgood and A. W. Heyer, Jr., A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118.

[8] Köhler, *op. cit.*, 1951, 239-240.

[9] Jaffe, *op. cit.*, 668-676.

[10] Köhler and Dinnerstein, *op. cit.*, 219; see also Köhler and D. A. Emery, Figural after-effects in the third dimension of visual space, this JOURNAL, 60, 1947, 159-201.

[11] John Krauskopf, Figural after-effects in auditory space, this JOURNAL, 67, 1954, 278-287, and Jack Nachmias, Figural after-effects in kinesthetic space, this JOURNAL, 66, 1953, 609-612.

Until a more inclusive physiological theory becomes available, it may be more profitable to interpret the phenomena reported in the present study as being related to certain other psychological concepts. In particular, the contrasting size of the visual stimulus-objects in Experiment I seems to be the essential condition under which there was an effect on the tactile and kinesthetic judgment. In Experiment II the visual stimulus was present only during the interpolated, satiating period, and no change in the judgment occurred. When, however, the 1-in. and 4-in. strips were seen contrasted with the 2-in. strip, there was a significant alteration. This suggests that the contrast between the visual objects resulted in a change in the general frame of reference, and this in turn influenced the kinesthetic judgment.

In terms of Helson's theory of an "adaptation-level," one might assume that the action of the visual stimuli is not limited to vision alone. These stimuli contribute to a common 'pool' of past and current effects of stimulation and thereby may produce alterations in all modalities.[12] Helson has shown that such 'pooling' of the effects of stimulation seems to determine subjective levels of reference to which comparative judgments may be related. Like Köhler's studies, Helson's work has been limited to influences within one modality. In contrast to a theory based on satiation, however, adaptation-level does provide a potential basis for the explanation of intermodal influences in perception. This broader application is possible since a specific physiological mechanism has not been invoked that is limited to a single sensory system.

SUMMARY

On the basis of the isomorphism embodied in Köhler's theory of satiation to explain figural after-effects, it was argued that after-effects should be specific to one modality. Twenty normal Ss were tested for kinesthetic figural after-effects according to the method of Köhler and Dinnerstein; but with the interpolated stimulus being identical with the standard. As expected, blindfolded controls showed no after-effects. Experimental Ss were exposed to additional visual stimuli, equal to the kinesthetic standard during test-periods and narrower or wider during satiation. These Ss showed a significant distortion in kinesthetic perception after seeing the wider stimuli. In a second experiment, Ss were exposed to visual stimuli only during the satiation period. Under these conditions no after-effects were demonstrated.

These results indicate that concurrent visual stimulation can induce a significant kinesthetic after-effect, but that this after-effect appears only when contrasting size-relationships exist between the visual stimuli. Such findings appear contrary to predictions made on the basis of the original satiational theory.

---

[12] Harry Helson, Adaptation-level as a basis for a quantitative theory of frames of reference, *Psychol. Rev.*, 55, 1948, 297-313; *Theoretical Foundations of Psychology*, 1951, 377-385, esp. 381.

THE NON-RECALL OF MATERIAL PRESENTED
DURING SLEEP

By William H. Emmons and Charles W. Simon, The Rand Corporation

A number of studies have been concerned with the possibility of learning during sleep, but their results have been inconclusive due to failure to determine satisfactorily whether S was asleep when the training material was presented.[1] In an earlier experiment, the present authors used the electroencephalogram (EEG) to monitor continuously the sleep states of their Ss. Questions and answers were read, one time each, at 5-min. intervals during an 8-hr. sleeping period, no matter what the S's sleep-state was at the time. It was found that sleep-states, as indicated by the EEG, correlated highly with recall. Since there was almost no measurable recall when material was presented during periods in which alpha frequencies (8-13 c.p.s.) were absent from the EEG record, learning during actual sleep did not seem possible nor learning during the deep drowsy state very practical.[2]

Could this outcome have been due to the failure to give repetitive training? In the present study, which was designed to supplement the first, the material was repeated as many times as possible, but only when the EEG indicated that S was asleep. It was hypothesized that material presented under these conditions would not be recalled.

METHOD

Subjects. Nine men were used as Ss. Their ages ranged from 18-31 yr. with a mean of 24.9. Their mean IQ, as measured by the Otis Self-Administering Test, Form D, was 112, with a range from 102-123. All of the Ss had participated in an earlier sleep-learning study and were familiar with the experimental environment. To facilitate the fine discriminations from the EEG records at the borderline between wakefulness and sleep, only Ss with persistent waking occipital alpha rhythms were used. A waking pattern typical of the Ss used is shown in Fig. 1. There is no reason to suspect that selection on the basis of the alpha rhythm would

have any effect on learning while awake or asleep.[3] A second group of 113 Ss (men, college students) was used to establish measures of preference for words being used as training material.

Apparatus. The Ss slept in clean, comfortable beds in three separate sound-proof, air-conditioned, electrically-shielded booths. The *EEG* electrodes were applied to the right occipital area, the vertex, and the left mastoid process. The leads were arranged to permit relatively free movement during sleep, and Ss suffered no dis-
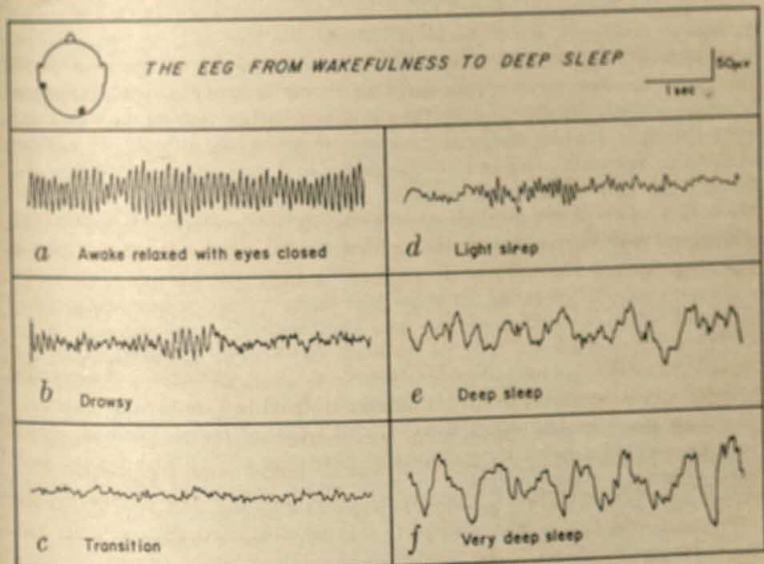
FIG. 1. ILLUSTRATIVE *EEG* RECORDS

comfort from the arrangement. Two monopolar recordings (right occipital and vertex) were made for each S, using a six-channel Offner electroencephalograph. A marker-pen mounted on the ink-writer automatically marked the exact time each word was presented. The learning material was recorded on magnetic tape and played through loudspeakers placed in the booths. The loudness-level was such that the material could be heard clearly, but not great enough to disturb a soundly sleeping S. A two-way intercommunication system allowed E to communicate with the Ss.

Criterion of sleep. Characteristic *EEG* patterns appearing during stages of waking and sleeping are illustrated in Fig. 1. These patterns correspond to those first described by Loomis, Harvey, and Hobart, and recently elaborated by the authors.[4]

[3] D. B. Lindsley, Psychological phenomena and the electroencephalogram, *EEG Clin. Neurophysiol.*, 4, 1952, 443-456.

[4] A. L. Loomis, E. N. Harvey, and G. A. Hobart, Cerebral states during sleep, as studied by human brain potentials, *J. exp. Psychol.*, 21, 1937, 127-144; Simon and Emmons, *op. cit.*, (in press).

The *EEG* of each *S* was continuously monitored, and the training material was played only when alpha frequencies were absent for 30 sec. prior to and during stimulation. As a further precaution, the training material was turned off if muscle potentials or movement artifacts, which tended to obscure cortical potentials, were observed in the record. When alpha frequencies were absent, as in *c, d,* and *e* of Fig. 1, *S* was considered asleep and the training material was played. These patterns have been related to the sleep-state. The *Ss* in one study typically signalled that they had 'floated' or 'drifted off' during the period when alpha was absent.[5] In another experiment, the *Ss* no longer turned off a tone as instructed when patterns such as *d* or *e* of Fig. 1 were observed.[6] In yet another investigation, *Ss* lost the ability to remember dropping an object shortly after alpha had disappeared.[7] Kleitman recently reported that no dreams occurred when patterns such as *d* or *e* were observed.[8] Lindsley reviewed the numerous studies on this subject and concluded that *Ss* were asleep when patterns similar to *d* were observed, although some investigators have correlated the onset of sleep with patterns such as *c*.[9] To give the benefit of the doubt to sleep-learning, the lighter level (*c,* where alpha has disappeared and 14-cycle spindle activity has not yet appeared) was selected as indicating the onset of sleep.

Training material. A list of 50 one-syllable nouns was given to a group of 113 untrained college students who were told that they were participating in a test of *ESP.* They were asked to pick the 10 words they believed that *E* had previously chosen by chance. This method made it possible to obtain an estimate of the probability of each word's being selected. By chance it would be expected that each word of the 50 would be selected 23 times. The 10 words of the 50 that were chosen for the training list had a mean frequency of selection of 22.9 with a range from 20-25. The order of presentation of the words on the training list during sequential repetitions was varied as the positions in a row of a homogeneous latin square.

Procedure. The *Ss* retired between 10-12 P.M. Before going to sleep, *Ss* were given the following instructions.

We are going to play a list of words to you while you sleep. We want you to learn them. If you should awaken during the night, say your name and booth number. If you should awaken and hear anything, say all of the words that you heard.

No attempt was made to play the test-material until 1 hr. after the final instructions had been given. If the *EEG* indicated that *S* was asleep and no alpha had been present for the previous 30 sec., a recording of irrelevant material was played; then, if the *S* still showed no signs of awakening, the test-list was played. It was immediately turned off if there was reason to doubt that *S* was still asleep. In some instances, *S* would tend to awaken each time the irrelevant material was played. In these cases the loudness was progressively diminished until it no longer awakened him, although it was never reduced below the previously determined level of intelligibility. The

[5] Hallowell Davis, P. A. Davis, A. L. Loomis, and E. N. Harvey, Human brain potentials during the onset of sleep, *J. Neurophysiol.,* 1, 1938, 24-28.
[6] Davis, Davis, Loomis, Harvey, and Hobart, Electrical reactions of the human brain to auditory stimulation during sleep, *ibid.,* 2, 1939, 500-514.
[7] Helen Blake, R. W. Gerard, and Nathaniel Kleitman, Factors influencing brain potentials during sleep, *J. Neurophysiol.,* 2, 1929, 48-60.
[8] Personal communication.
[9] Lindsley, *op. cit.,* 443-456.

material was played to each S as many times as possible. If the S tended to awaken easily, the material was turned off long enough to allow him to go into a deep sleep. The number of times each S awoke and whatever he said at these times were recorded. Due to the necessity of turning off the material when the S showed signs of awakening, it was not possible to present each word an equal number of times. It seemed safe, however, to assume that the procedure was a random one, and an examination of the number of times each word was presented supported this assumption.

Upon awakening in the morning and before leaving his booth, each S was asked to write down any words he had heard or thought he had heard during the night. The papers were collected and the Ss were then given the same list of 50 nouns which was used to select the 10 training words. The Ss were instructed to read every word on the list and to select the 10 words that were played during the night. Upon completing their selections, the Ss were allowed to dress and wash, after which they were given the list of 50 words a second time and again instructed to select the 10 words played during the night. They were instructed that they could either choose the same words they selected the first time or they could change their selections, whichever they wished. The Ss then returned to their booths and the list was played through once at the minimal loudness at which it had been played during the night. The Ss were asked to repeat each word as it was played. Every word was correctly repeated by all Ss.

RESULTS

The results of this experiment give no evidence that auditory material could be recalled after being presented a number of times during sleep. From Table I (A), it is clear that the trained experimental group failed to

TABLE I

MEAN PERCENTAGE OF TRAINING WORDS CORRECTLY SELECTED BY THE CONTROL GROUP AND BY THE EXPERIMENTAL GROUP ON THE FIRST TEST

	Mean % correct	SD	N	t*	df	P
CONTROL	20.3	11.8	113	—	—	—
EXPERIMENTAL						
(A) All training items	24.4	8.8	9	1.32	120	0.20
(B) Items with no alpha following within 10 sec.	12.0	19.1	9	-1.28	120	0.25
(C) Items with alpha following within 10 sec.	33.4	17.8	9	2.18	120	0.04
(D) Items with alpha following, but items reported heard removed	23.9	18.2	9	0.58	120	0.50

* When the data from all Ss were combined and Chi-square tests were made, the same conclusions were reached as those drawn from the t-tests.

select significantly more correct words than the untrained control group on the first test. Results for the second test were so similar to those for the first that there is no need to present them here.

A further analysis was made with the *S*s acting as their own controls. From the list of 50 words, 10 new ones were selected which had been chosen by the control group with the same frequency as the words in the training list. It was found that the experimental *S*s selected on the average the same percentage of words from the untrained list as they did from the trained list. These results further negate the hypothesis of sleep-learning.[10]

Nevertheless, with continuous *EEG* monitoring while the stimuli were being presented, it was possible to partial out that material which had been presented just prior to the reappearance of alpha frequencies in the *EEG*, when the input was turned off. In an earlier experiment, a portion of the material presented under these conditions—representing a deep drowsy state—were reported heard and could be recalled.[11] When only this material is considered, Table I (C) shows that the experimental group did select significantly more training words than the control group. There was no significant difference between the two groups for the material remaining after the items presented just prior to alpha were removed (B).[12]

Other sleep-learning investigators have eliminated *S*s who reported hearing the training material. In this experiment, four *S*s reported hearing a total of eight words at the time of presentation. Only two of these words were recalled correctly the next day, five were correctly selected on both tests, and three were never correctly selected. When the words reported heard were removed from those items occurring just prior to the occurrence of alpha, the difference between the experimental and control groups was no longer significant (Table I, D). It would seem, therefore, that *S* must be able to report hearing the material before he is able to recall it.

These negative findings tend to refute the argument that the failure to demonstrate sleep-learning in the first experiment was due to the lack of repetition of the material. In the present study, the mean number of repetitions for all *S*s was 46.3 with a range from 16.0-81.6. Our results are in agreement with those of Stampfl and of Coyne who found no relationship between number of repetitions and amount of sleep-learning.[13]

---

[10] The recall of a group of 37 control *S*s improved 47% after they heard a list of 10 words only once, with no instructions to learn and an intervening half-hour period filled with other activity.

[11] Simon and Emmon, *op. cit.,* (in press).

[12] It is interesting to note that, of the 15 items consistently selected correctly by the 9 *S*s on the first and second tests, 14 had been presented just before the appearance of alpha. Of the correct items with respect to which the *S*s changed their minds from the first to the second tests, only 3 of 7 were presented just before the appearance of alpha.

[13] T. G. Stampfl, The effect of frequency of repetition on the retention of auditory material presented during sleep, Master's thesis, Loyola University, Chicago, 1953, 25-31; M. G. Coyne, Some problems and parameters of sleep learning, Honors' thesis, Wesleyan College, Connecticut, 1953.

Stampfl has suggested that negative findings in sleep-learning studies may be due to the fact that the Ss were in too deep a level of sleep during the presentation of the material.[14] In the present study an analysis was made of the *EEG* records during the presentation periods in accordance with the sleep-level categories shown in Fig. 1. It was found that no material was presented while the Ss were wide awake, 1.2% of the material was presented in the deep drowsy state, 15.9% was presented in the transition state where sleep begins, 57.3% occurred in light sleep, 22.2% in deep sleep, and 3.4% in very deep sleep. It is unlikely that the negative results in this study are due to too deep a sleep-level at the time of stimulus-presentation, since 74.4% of the repetitions occurred at the borderline and at the lighter levels of sleep.

The importance of continuous *EEG* monitoring while presenting the test-material in sleep-learning studies is shown by the fact that material could be played on the average of only 2.4 min. without S's showing some signs of awakening. Although this interval may increase after an additional period of adjustment, the problem cannot be overlooked by future experimenters, for the Ss in this study were already experienced, having previously spent one or more nights in the laboratory under similar conditions.

SUMMARY

A list of 10 one-syllable nouns was repeated as many times as possible to 9 Ss during an 8-hr. sleeping period. A continuous *EEG* recording during the presentation of the training-material was used to determine the sleep-level at that time, and the material was turned off as soon as cyclical activity within the alpha range was observed. The experimental Ss did not do significantly better than the controls in selecting the words on the training list from a list of 50 words. Nor did they choose the training words any more frequently than they chose an equivalent group of control words. There was some indication that words presented during a period of deep drowsiness can be retained, but this tendency was significant only when the S was also able to give an immediate response to the material being presented. The effects of sleep-level and the importance of continuous *EEG* monitoring while presenting the training-material are discussed in their relation to recall. It it concluded that material presented a number of times during sleep (using an *EEG* criterion) cannot subsequently be recalled.

---

[14] Stampfl, *op. cit.,* 32.

RECOGNITIVE THRESHOLDS FOR WORDS AS A FUNCTION OF SET AND SIMILARITY

By Sherman Ross, Matthew Yarczower, and Gerald M. Williams,
University of Maryland

In a recent study Neisser investigated the distinction between perceptual process and verbal response.[1] He concluded that S's 'set' in tachistoscopic experiments "facilitates the perception of specific visual patterns."[2] The specific visual patterns used were words which appeared in a preliminary observation and were presented again in the test-period. The results of Neisser's study indicated that if items were presented to the S in a pre-test and again in a test, the recognitive thresholds were lower for these repeated words than for their homonyms. The structural characteristics of the words used were not controlled, as Neisser pointed out.

The present investigation, a modification of Neisser's study, was carried out to determine the relationship between preparatory set and the structural characteristics of the words to the recognitive threshold. If Ss are presented with a series of stimulus-words and are instructed that this presentation may help in recognizing other stimuli to be presented later, can we say that this 'set' only affects the recognitive threshold when the same words are repeated? Does not the 'set' affect recognition of other words which may bear a certain relationship to the original ones? There is some evidence that the physical characteristics of the stimulus affect recognitive threlolds.[3] Two major reviews related to this topic have recently appeared in the literature.[4]

The theory is advanced that the recognitive thresholds in tachistoscopic experiments are affected by the relationship between the 'set' for the stimulus-words and the test-words. The relationship investigated in this experiment is one of similarity of construction. We advance the following hypotheses: (1) that the preparatory 'set' affects not only the stimulus-words repeated but also words bearing a structural relationship to the words initially exposed—specifically that 'set' will facilitate the recognition of words that are shown in a preliminary trial and repeated in the test-trial; (2) that it facilitates the recognition of words which are similar to the 'set' words; and (3) that its effect will decrease with increasing *dissimilarity* between the 'set' and the test-words. We assumed that the underlying continuum of structural similarity was in terms of perceptually similar units. On this basis, three groups of test-words were so chosen that these groups represented different points on the similiarity continuum: similar, intermediate, and dissimilar.

[1] Ulric Neisser, An experimental distinction between perceptual process and verbal response, *J. exp. Psychol.,* 47, 1954, 399-402.
[2] Neissed, *op. cit.,* 402.
[3] Elliott McGinnies, P. B. Comer, and O. L. Lacey, Visual-recognition thresholds as a function of word length and word frequency, *J. exp. Psychol.,* 44, 1952, 65-69.
[4] L. J. Postman, The experimental analysis of motivational factors in perception, In *Current Theory and Research in Motivation,* University of Nebraska Symposium, 1953, 59-101; F. H. Allport, *Theories of Perception and the Concept of Structure,* 1955, 337-361.

Materials and Procedure. The words were presented to each *S* individually, using a modified Dodge tachistoscope. Illumination was provided by small fluorescent lamps, which turned on and off instantaneously. Stimulus-words were typed in capital letters, parallel with the long side of an $8\frac{1}{2} \times 11$ in. sheet of paper. The words were centered on the sheet. These sheets were then fastened to cardboard slides, which moved easily through the back of the tachistoscope.

The *S*s for the experiment were 45 undergraduate students (26 men and 19 women) at the University of Maryland. The *S*s were assigned to five experimental groups on a random basis. Ten 'set' words, peculiar to his assigned group, were presented tachistoscopically to each *S*. The words were then repeated. The order was randomized for each presentation and the words were shown individually at a suprathreshold value, arbitrarily selected at 0.60 sec. *S* was asked to write each word,

TABLE I

WORDS USED IN PRELIMINARY (SET) AND TEST-PERIODS

Group	'Set' Words	Test-Words
(1) Practice	be, but, the, for, or, wail, phrase, one, colonel, islet	be, but, the, for, or
(2) Similar	be, but, the, for, or, wail, colonel, phrase, one, islet	bee, butt, thee, fore, ore
(3) Intermediate	break, bail, board, fare, allowed, for, the, but, or, be	brake, bale, bored, fair, aloud
(4) Dissimilar	phrase, islet, colonel, one, wail, board, allowed, bail, break, fare	frays, eyelet, kernel, won, whale
(5) Control	wail, but, or, colonel, the, phrase, one, for, be, islet	choose, trust, sort, fort, nose

as it was presented, on the back of his data-sheet. He was told that the words were presented in this way to familiarize him with the words and with the task. *S* was instructed to write the words in order to aid him in remembering them since this might help him on his task to follow.

The test-words of the five experimental groups were given as follows: Group 1 (*Practice*) received five of the words appearing in the 'set' list; Group 2 (*Similar*) received homonyms of five of the 'set' words which were characterized by having only one more letter than the corresponding 'set' word; Group 3 (*Intermediate*) received five homonyms of the 'set' words, which were judged to be of intermediate similarity to the corresponding 'set' words; Group 4 (*Dissimilar*) received five homonyms of the 'set' words which were judged to be dissimilar to the corresponding 'set' words; and Group 5 (*Control*) received five new words, which had no predetermined connection with the 'set' words. The test-words were presented to each *S* in random order. 'Set' and test-words are shown in Table I. Thresholds were determined by a modified Method of Limits with the first presentation of each word at 0.04 sec., followed by increments of 0.04 sec. The criterion of recognition was three consecutive correct responses as written by *S*. The exposure-time of the first correct response in the series was taken as the threshold-value. *S* was told that the test-word would be presented very quickly, but would be repeated until *E* designated that the trial was completed. He was also told that he should not necessarily interpret this

to mean that he had not seen and recorded the correct word. S was instructed to record on the data-sheet exactly what he thought he had seen after each presentation regardless of how foolish the response appeared. Frequency of usage was not a specific control. In every case, with two exceptions, the test-words (similar, intermediate, dissimilar) have a lower frequency-count than do their corresponding members in the 'set' list.

Results and discussion. Table II shows the mean recognition thresholds for all test-words for each experimental group. The means were 0.102, 0.319, 0.196, 0.307, and 0.191 sec. for the Practice, Similar, Intermediate, Dissimilar, and Control groups respectively. A *t*-test was made for each of the 10 mean differences among the groups.

TABLE II
MEAN RECOGNITIVE THRESHOLD AND VARIANCE FOR ALL TEST-WORDS FOR EVERY GROUP

Group	Threshold (in sec.)	Variance
(1) Practice	.102	.005
(2) Similar	.319	.011
(3) Intermediate	.196	.008
(4) Dissimilar	.307	.011
(5) Control	.191	.022

These results are shown in Table III. In several cases, a preliminary test indicated that the assumption of equal variances was suspect. Hence an adjustment was made in terms of the degrees of freedom.[5] There is no significant difference between the

TABLE III
THE *t*-VALUES FOR THE MEAN DIFFERENCES AMONG THE GROUPS

Group	(2)	(3)	(4)	(5)
(1)	5.229†	2.706*	0.246	3.396†
(2)		2.509†	4.944†	3.230†
(3)			2.444†	0.150
(4)				3.083*

* Significant beyond the 5-% level of confidence.
† Significant beyond the 1-% level of confidence.

means of the Intermediate and Control Groups, nor between those of the Similar and Dissimilar Groups. All other comparisons exhibit statistically significant differences between the means.

It was expected that if the preparatory 'set' facilitated recognition of a specific visual pattern, it would facilitate recognition of visual patterns similar to the specific pattern. The hypothesis stated above was drawn from the theory of identical elements. It asserts that training in one activity facilitates another activity only insofar as the two have elements in common. It was assumed that the elements were *perceptually* identical elements. The view that principles of learning can be applied to perceptual problems is evident in the following statement: "It appears that there are important

---

[5] H. M. Walker and J. Lev, *Statistical Inference,* 1953, 157-158.

communalities between the principles of verbal learning and the principles of perceptual recognition."[6] Solomon and Postman state that "the phenomena of word recognition can play a strategic rôle in the rapproachement of theories of perception and verbal learning."[7]

In terms of our original hypotheses we find that the preparatory set did facilitate the recognition of specific visual patterns. Group 1 (Practice) had a significantly lower threshold than did Group 5 (Control). Hence, Neisser's results are confirmed when tested under different experimental conditions. The recognitive thresholds as a function of structural similarity did not yield a monotonic function as we hypothesized. The means for Group 2 (Similar) and Group 4 (Dissimilar) were significantly greater than the Group 3 (Intermediate) mean. Both groups (Similar, Dissimilar) had significantly higher thresholds than the Control Group. Hence, there was a detrimental effect in the recognition of the test-words for the two groups. The difference in recognitive thresholds between the Intermediate and Control Groups was not statistically significant. An analysis of Ss' responses was made in terms of the number of interference responses in each of the three groups (Similar, Intermediate, and Dissimilar). An interference response was defined as the reporting of the paired member of the test-word when the test-word was presented to the S, $e.g.$ when 'brake' is presented, in the test-trial, the S responds with 'break.' Tests of the significance of the difference between proportions were run for all possible comparisons. The there comparisons were all statistically significant beyond the 1% level. The group giving the greatest number of interfering responses was the Intermediate Group. The smallest number of interfering responses was given by the Dissimilar Group. The Similar Group was between the two. These results would indicate that the greater recognitive thresholds could not be attributed to interference responses, as defined.

Examination of the recognitive thresholds as a function of word frequency yielded a rank-order correlation coefficient of 0.597. When one examines the recognitive thresholds of words in the three groups (Similar, Intermediate, Dissimilar) that have equivalent frequency counts the same results are obtained. An examination of the words used indicates that 80% of the words in the Intermediate Group had the same number of letters in the test-word as in the 'set' word. In the Dissimilar Group 20% of the test-words had the same number of letters as the 'set' words. The Similar Group had none, since to each 'set' word one letter was added to make up the test-word. It is quite possible that the percentage of agreement in length of word between the test and 'set' words played a significant rôle in the reporting of interference responses and consequently in the recognitive thresholds.

In a recent paper by Bricker and Chapanis, the question was raised as to whether or not incorrectly perceived tachistoscopic stimuli convey some information.[8] The authors concluded, "Even when S gives an incorrect verbal response to a stimulus, the stimulus may still have conveyed some useful information."[9] We would like to

[6] Postman, $op. cit.,$ 76.
[7] R. L. Solomon and Leo Postman, Frequency of usage as a determinant of recognition thresholds for words, $J. exp. Psychol.,$ 43, 1952, 200.
[8] P. D. Bricker and Alphonse Chapanis, Do incorrectly perceived tachistoscopic stimuli convey some information?, $Psychol. Rev.,$ 60, 1953, 181-188.
[9] Bricker and Chapanis, $op. cit.,$ 185.

raise the question, in light of our results, as to the possibility of a differential amount of information being conveyed, as a function of the relationship between the stimulus presented and the incorrect verbal response. Assume that the word 'be' is shown in the preliminary session. The word 'bee' is then presented, tachistoscopically at relatively short exposure-times in the test-period. Let us say that the S responds with 'be,' an incorrect interference response in the test-period. Does this response increase the probability of the correct response 'bee' being made? It has been shown that characteristics of an initial learning task do transfer to subsequent learning tasks on the basis of the homophony of the learning material.[10] The question raised here must be answered by future experimentation.

In light of the previous observations, the hypothesis is ventured that the continuum underlying the perceptually similar elements is a complex one with length of word and number of *identical elements* contributing a significant portion of the variance of the recognitive thresholds for the test-words. The effect of this similarity continuum on recognitive thresholds in tachistoscopic experiments is counfounded by interference effects and by possible changes in response probability.

---

[10] T. G. Andrews, Solomon Shapiro, and C. N. Cofer, Transfer and generalization of the inhibitory potential developed in rote serial learning, this JOURNAL, 67, 1954, 453-463.

THE RÔLE OF CONTENT IN BINOCULAR RESOLUTION

By Edward Engel, Princeton University

In studies of binocular rivalry, the problem of predominance has been treated exclusively in its relation to the formal properties of the stimuli presented.[1] Little if any attention has been given to the content of the discrepant patterns as a possible source of influence upon the binocular outcome. Helmholtz came closest to considering the influence of content when he proposed an attentional theory of binocular rivalry in which the interest-character of the objects viewed was accorded a rôle of some consequence.[2] Beyond this, the literature is quite free of reference to content, and the omission is reflected in the use of such abstract stimuli as circles, squares, and colored patches.

The experiment at hand represents an attempt to study monocular predominance as a function of the orientation of familiar objects; namely, human faces. A stereoscopic arrangement was used which presented a face in the usual vertical orientation to one eye of O, and an inverted face to the other eye. The two monocular faces stimulated corresponding areas of the two retinas and therefore converged upon the same portion of the binocular field. The hypothesis tested was that the face in normal, functional position with respect to O would be favored—that a reliable tendency for the upright face to predominate would appear.

METHOD

Apparatus. The basic apparatus was a modified stereoscope enclosed in a light-tight box supported by four small legs. When this apparatus was placed on a standard sized table the height of the viewing-piece was a comfortable one for the seated O.

The interior of the apparatus contained the shaft of a conventional stereoscope suspended 1 in. from the floor of the box. A stereogram-holder was attached to the shaft, and a pin which allowed the holder to be moved along the shaft was fastened to the bottom of the holder. The floor of the box had a narrow opening, ⅓ in. in width, which ran the length of the shaft and directly below it. The holder-pin protruded through this opening 2 in. below the floor of the box, enabling O to adjust the stereogram.

Running along the shaft, and above it, was a length of opaque black cloth equiva-

[1] See M. D. Vernon (*A Further Study of Visual Perception*, 1952, 204-205) and R. S. Woodworth (*Experimental Psychology*, 1938, 573, 659.) for surveys of the literature on binocular rivalry.

[2] H. von Helmholtz, *Treatise on Physiological Optics.* (J .P. C. Southall, ed.) 3, 1925, 498-500.

lent in height to the stereogram. On one end, it was attached to a vertical roller at the front end of the shaft between the two lenses. On the other end, it was attached to the center of the stereogram-holder. By means of a spring-arrangement, when the holder was moved away from *O*, the roller released an additional length of cloth, and when the holder was moved toward *O*, the slack in the cloth was taken up. This arrangement so separated the fields of the left and right eyes that neither eye could view any portion of the inappropriate field.

In the front right and left hand corners of the interior, two 7-w. lamps were suspended which illuminated the stereogram. Two variacs outside the box regulated the flow of current to these lamps. The variacs were set at 70 v. The usual fluctuations, of course, occurred, but they were small, and equivalent for the two lamps.

Two stereograms were employed. They were constructed by stapling together two squares of cardboard, one of which served as the front of the stereogram, the other as the back. Two squares, each measuring 1¾ in., were cut out of the front board, and the photo of a man's face was inserted into each square, one of the faces
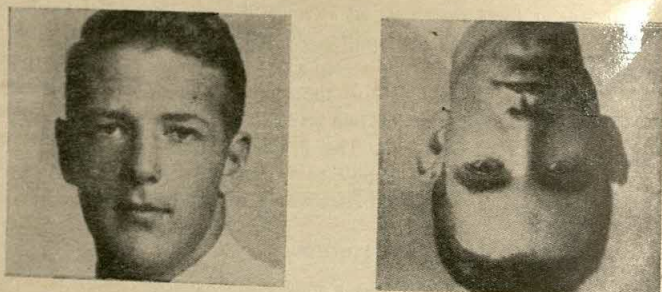
FIG. 1. THE TWO FACES AS THEY APPEARED ON STEREOGRAM *A*

inverted. The distance from the center of one cut-out to the other was 3 in. The stereograms were of standard size, measuring 7 × 3⅝ in. The faces of Stereogram A are shown in Fig. 1. Stereogram B was the same as A except that the faces were, in the orientation shown, interchanged. The two stereograms were used also in the inverted position, thus reversing the orientation of the faces of each and giving a total of four presentations.

Observers. Os were 12 men aged 18-65 yr. They had no history of visual defects of any kind, never having worn corrective glasses nor ever having been advised to do so. In general, the Os had 20/20 vision or better. The only exceptions were the two oldest Os, whose vision was found to be equivalent for the two eyes but somewhat less effective than that of the others. In the case of all Os, wherever any doubt existed as to the approximate equivalence of the two eyes, an eye-test was given.

Procedure. O was seated in position and a stereogram having a small circle as one monocular component, and a large circle as the other, was inserted in the stereoscope. He was then instructed to manipulate the adjusting pin until the small circle was centered within the larger circle. When O reported that the circle was centered, he was informed that it would no longer be necessary to make any further adjustments in subsequent observations. O then observed Stereograms A and B, each in two

orientations. First, *A* upright (Fig. 1); second, *B* inverted; third, *A* inverted; fourth, *B* upright. For each condition, *O* was instructed, "Place your head in the viewing position and close your eyes." The light-switch was then turned on and *O* instructed, "Open your eyes and describe what you see." If, after 1 min., *O*'s report did not clearly contain information as to the relative dominance of the two faces, the following question or its equivalent was asked: "Does either face predominate or are they equally predominant?" This question was deemed unnecessary in the following cases: (1) where *O* stated of his own accord that one or the other was more prominent or that they were equally so, which happened most often after the first presenta-

TABLE I

RESPONSE-DISTRIBUTIONS FOR EACH *O*

O	Upright predominant	Equal	Inverted predominant
1	4	0	0
2	4	0	0
3	2	1	1
4	4	0	0
5	4	0	0
6	2	2	0
7	4	0	0
8	2	1	1
9	4	0	0
10	4	0	1
11	3	0	0
12	4	0	0
Total	41	4	3

tion, *O* having already been asked the question; (2) where *O* stated that he saw the upright or the inverted face clearly while the other was hardly evident; (3) where *O* stated that one of the faces always was present while he got only fleeting glimpses of the other; (4) where *O* described a single face and gave no sign that he was aware of the second.

RESULTS

Of the 48 responses given by the group as a whole, the upright face predominated in 41, the inverted in 3, and there was no clear predominance in the remaining 4. Every one of the *O*s reported more upright pictures as predominant than inverted pictures, an outcome which is significant at a level of confidence well beyond 1% ($P = 0.5^{12}$). There was no consistent difference between either with regard to eye or to face. The pattern presented to the left eye dominated 22 times, the pattern presented to the right eye dominated 22 times, and there was no difference in the remaining 4 cases. One of the two faces dominated 21 times (the left face of Fig. 1), the other dominated 23 times, and there was no difference in the remaining 4 cases.

The following verbatim transcript of the report of one of the *O*s is representative:

First presentation. I see a boy—a boy's face, but there is something else there that I can't see at the moment. Something interfering with this. Ah—now wait a minute. There's something upside-down. I can't see it too well. I would guess it was some-body's face upside-down but I can't see it. I can see the nose at the top of the boy's

face. It comes and goes, but I can't see much of it. I can see his hair; oh wait a minute—he's coming in now. No, he went again. I can see his hair if I look down and it looks like a beard on the young man's face. He goes away again. If I look up I can see his mouth and his nose a bit. The eyes must correspond. That must be the secret of this. But the eyes of the two if I look up, no I can't really, the best I can get in the way of seeing two is to look up and then I can see them both with beards as it were. [E: Which one dominated in the picture?] The upright one. All the time or 99.9% of the time I could see him completely.

Second presentation. Another boy, but there is something definitely behind there, but I can see him pretty well clear at the moment. Now the other one is coming through. Oh—now I can see it upside-down, just for a fleeting instant—yes it's always the same boy, *i.e.* the photograph upside-down. I get the fleeting impression of a gorilla in the middle of this. Now I can see the one right-way-up with just a few bits of the other. The other boy is coming back— not really though. Ah, yes, he's there now, just for a second though. Maybe I can get him back if I work hard on him. Yeh, there he is. It's not the same boy though, is it? [E: Which one was dominant?] The vertical one; the right-way-up, but not as much this time as last time.

Third presentation. The same boy is back. Again there's somebody else there. And ah—oh I can see him this time—nearly. I got him for a second and he went away a second after. Now there is a strange sort of a mixture there—yes, I can get—I got the second boy now, the upside-down face. Why, you get a horrible view when they both come—when they clash. There's the vertical, the straight-up-and-down man again. Yes, I can get the upside-down man. I'm getting better at this. It's mainly the vertical man, but if I work very hard I can get the upside-down man on his own.

Fourth presentation. I can see the—oh the upside-down man has come, and almost first this time and I can see the other man on his own now. Now the upside-down man is pretty clear but those eyes are bothering me this time more than they did last. Yeh, the upside-down man is there in toto now. The other man has come back. I can't [pause] the right-way-up chap is spoiling it. Most of the time I can see him in the background. I don't seem to get this clearing away that happened last. I can see the [pause] never does the right-way-up man fade out of the picture entirely. His hair is always there. His eyes particularly are always there. The upside-down [pause] I could see practically all of him. I could almost describe him but the other fellow is always there. He's dominant alright—he tends to be there. In fact, I can't get rid of him—not entirely.

DISCUSSION

The results obtained confirm the prediction that an upright face has a definite and reliable advantage with regard to predominance in the binocular field. Of primary interest is the fact that the effect is visual in nature and not simply a judgment made upon two equally well-represented figures. The upright figure tends in general to be favored with respect to clarity, definiteness, coherence, and continuity of parts. The inverted figure tends on the whole to be more disjointed, more obliterated, and generally less well and less often represented. These effects have a distinct sensory character and cannot be considered merely 'interpretive' in nature.

The predominance of the upright figure is exhibited in two distinguishable ways. There is, first of all, a tendency for the upright face to predominate when the two figures are simultaneously perceived. In this case, the tendency is for the upright face to emerge as an organized figure whose appearance is marred somewhat by isolated and partial details and contours of the inverted face. These details are perceived as conflicting with, and extraneous to, the dominant impression. To a lesser extent, the

reverse occurs with the inverted face emerging as figure and the upright as contributing extraneous detail. Secondly, there is a tendency for the upright figure to predominate where the two figures are perceived alternately in the field (binocular rivalry). In this case, the upright figure tends to remain longer in awareness during the successive phases of alternation.

What is here clearly in evidence is that the two figures, equal in physical stimulus-properties, differ in their influence on sensory organization in a manner which depends on their spatial orientation. Furthermore, the particular orientation wihch is favored is that which has been more frequently encountered in the past. This, coupled with the fact that orientation, at least in this instance, cannot be readily construed as inherent in the forms, *per se,* suggests that these earlier encounters exercise a definite influence on subsequent organization. Precisely how these effects are mediated remains to be determined.

A theory of perception which holds that sensory organization is wholly prior to, and independent of, content would appear to be contradicted by these results. In espousing such a theory, Köhler noted that "given specific entities, with their shapes" might acquire meanings. "But when this happens," he wrote, "these entities are given first, and the meanings attach themselves to such shaped things later."[3] In discussing Gottschaldt's works, Köhler proposed that "those who claim that past experience has an automatic effect upon subsequent perception will have to support their theory by experiments of their own." The present study seems to provide such support.

---

[3] Wolfgang Köhler, *Gestalt Psychology,* 1929, 192.

AN EVALUATION OF A METHOD FOR DEVELOPING
RATIO-SCALES

By Trygg Engen, Brown University

This study is an experimental investigation of some aspects of Metfessel's constant-sum method for developing ratio-scales.[1] Metfessel suggests that O be instructed to make a "direct estimation of the ratio between the psychological magnitudes corresponding to two stimuli." O is allowed some convenient number of points, usually 100, and has to divide these points between the two objects so as to express the ratio between some aspect of the experience corresponding to the objects. A perceived ratio of 3 to 2 would thus be reflected by the assignment of 60 and 40 points. Metfessel stated that the judgments will then directly express psychologically meaningful ratios because the judgments represent perceived quantitative relations among the objects.

Comrey, the first investigator to employ the method, applied it to judgments of lines, and Nelson continued the exploration by the method of lifted-weights.[2] Both of them developed scales which closely approximated measurement of their stimulus-materials. Guilford and Dingman compared the constant-sum method with the fractionation-method and found that the two methods yielded completely concordant results for paired comparisons of weights that were judged to be in the ratio of 2 to 1.[3]

The purpose of the present experiment was to investigate the effects on the psychological scale that are produced by variations (a) in the size of the range over which the stimuli are distributed and (b) in the absolute magnitude of the stimuli.

Procedure. The stimuli consisted of 11 lines of various lengths, and for purposes of identification each line was designated by a letter. The line lengths in inches were as follows: A, $\frac{1}{2}$; B, $\frac{2}{3}$; C, 1; D, $1\frac{1}{2}$; E, 2; F, 3; G, $4\frac{1}{2}$; H, 6; I, 9; J, $13\frac{1}{2}$; and K, 18. All lines were $\frac{1}{4}$ in. thick. The 11 lines were arranged into four ranges, I, II, III, IV, as follows:

I. A	B	C	D	E	F	G	H	I	J	K
II. A	B	C	D	E						
III.			D	E	F	G	H			
IV.						G	H	I	J	K

The answer to the first question, which was concerned with the effect of the size of the range, involved comparisons of the scale-values obtained from the same

stimuli when they appeared in the extended range (Range I, A–K) with those obtained from the shorter ranges (Ranges II–IV). The stimulus-ratio of A to K was 1 to 36, while that within the ends of the shorter ranges was always 1 to 4.

The second question had to do with the effect of the absolute magnitude of the stimuli. For this problem Ranges II, III, and IV were considered as sections of Range I, and each of these ranges was presented independently. The short ranges were identical in terms of ratios represented and differed only with respect to actual length of line; hence, if different scale-values were obtained for these ranges, they could be attributed to the differences in the magnitude of the stimuli.

Since this study was a test of a method, it was not concerned with sampling considerations as such, and only 5 Os (graduate students in psychology) were used. O made one judgment on each possible stimulus-pair of all four ranges in each of

TABLE I

Physical and Psychological Scale-Values for Range I (Full Range) and for Combined Ranges II, III, and IV (Partial Ranges)

Physical length	Range I	Ranges II–IV Combined
1.000	1.000	1.000
1.333	1.402	1.437
2.000	2.053	2.222
3.000	2.932	3.385
4.000	3.772	4.399
6.000	5.542	6.752
9.000	8.075	10.227
12.000	10.588	13.667
18.000	15.751	20.797
27.000	22.612	31.125
36.000	29.559	41.755

six sessions. A session lasted about 90 min. and the time between sessions was about 48 hr.

The stimulus-pairs within the four ranges were presented in four different random orders with the one stipulation that no line was to occur in two consecutive pairs. Differential practice effects on pairs were controlled by counter-balancing their presentation within each range. The order of presentation of ranges was also counterbalanced to reduce series effects.

Photographic slides were made of all possible pairs of the 11 lines which were projected from the rear on a translucent screen that separated two darkrooms. The projected lines were white and of the stated lengths. They were horizontal with one line 6 in. above the other, but not aligned at the ends. The image on the screen was 7 ft. from O and at eye-level. E gave O instructions and recorded his responses through a two-way intercommunication system.

Results. The raw data consist of the total number of points assigned to each line in comparison with every other line. Inter-pair ratios and scale-values were determined by procedures suggested by Comrey.[4] Tables I and II show the resulting psychological scales, related to physical scales based on physical measurement of the lines.[5] Note that the low end of each scale has been assigned the value of 1.00.[6]

---

[4] Comrey, *op. cit.*, 317-321.
[5] Essentially the same results were obtained from a shorter repetition of the experiment, using 15 graduate students for two sessions each. The detailed results

The comparison between the long and the short ranges shows that the psychological scale-values for the long range (Range I) are smaller than the physical scale-values, whereas the scale values obtained with the short ranges (Ranges II–IV) are on the average greater than the physical scale-values. That is, a given stimulus is assigned a scale-value that is smaller or larger than its physical one, depending on

TABLE II

Physical and Psychological Scale-Values for Short Ranges II, III, and IV

Physical length	Range II (low)	Range III (medium)	Range IV (high)
1.000	1.000	1.000	1.000
1.333	1.437	1.303	1.345
2.000	2.222	2.001	2.047
3.000	3.385	3.030	3.564
4.000	4.385	4.022	4.?0

whether it was judged in the context of a long or short range of stimuli, respectively. Closer scrutiny of the ratio-judgments from which these scale-values were developed indicated that nearly all ratios reported in Range I were less extreme than the physical values; i.e. a physical ratio of 35-65 might be estimated as 40-60. As a result, there is an increasing negative discrepancy between the psychological and physical scale-values due to the fact that the scale is determined by successive multiplication of average ratio-judgments. Thus, the greater the magnitude of the stimulus, the greater is the discrepancy. This is a function of the method, since the discrepancies between physical and psychological assignments, that is, individual ratios, were found on the average to be no greater for the long range than for the short ranges. They did tend to be in opposite direction.

Table II shows that in Range II, which contained relatively small stimuli, the scale-values correspond to the physical scale less closely than in Ranges III and IV, consisting of the medium and large stimuli. This finding indicates that the magnitude of the stimuli had an effect on the judgments.

Since Ranges II, III, and IV were overlapping sections of Range I, they were combined by using the average of the values for D to E to go from Range II to III, and G to H from Range III to IV. Table I shows that this treatment of the data again leads to greater discrepancy in approximating the physical scale-values, i.e. the greater the size of the scale-value the greater the discrepancy. In this case small positive errors were exaggerated by successive multiplication in developing scale-values.

Discussion. Certain implications of the results require more detailed consideration. Increasing the range of stimuli presented for judgment in a series seemed to constrict the resulting scale; identical pairs yielded consistently different ratio-judgments in different contexts, specifically, the long and short ranges. The differences observed between the short ranges, which were identical except for the magnitude of the stimuli,

---

are available in MicA54-3186, University Microfilms, Ann Arbor, Michigan, and will not be reported here.

[6] Since the shortest line in each range was used as the unit, there is complete correspondence between physical and obtained scale-values at the value of one. Likewise, since Ranges II, III, and IV were identical, except for the absolute magnitude of the stimuli, the physical scales for these ranges are also identical.

show the effect of context in a different way. These findings suggest that there is an effect of the adaptation-level on these ratio-judgments.[7] The results indicate that when at least some of the differences between members of a pair were relatively large, O had a tendency to underestimate the differences, that is, assigning 45-55, for example, rather than 40-60. When on the other hand the differences between the stimuli of a pair were generally small, such as in the short ranges, O had a tendency to overestimate the difference.

The computational method also contributes to the apparent difference in results between the long and short ranges. Scale-values are determined by successive multiplication of average ratio-judgments, and this procedure has the effect of exaggerating whatever discrepancies may be present in these averages. This resulted in a relatively large discrepancy between the scale values and the physical values for the long range.

A limitation, due to the method of reporting the judgments, appeared in our results. This is that very large stimulus-differences, i.e. representing ratios greater than about 10:1, not only seem to represent more difficult discriminations but, even if such great ratios could be judged accurately, it is not possible to report the judgments as accurately as judgments of smaller stimulus-differences. For instance, it is impossible to express a ratio that the O has judged to be not quite 95-5, because while point assignments of 95-5 to the two stimuli represent ratios of 19:1, 96-4 and 94-6 represent ratios of 24:1 and 15⅔:1 respectively. This change in the sensitivity over the 100-point scale would have the tendency to make O's judgments of relatively large differences not only less accurate but less variable as well. It was found that an analysis of variance of judgments involving quite similar (i.e. ratios 1½:1 and smaller) and very discrepant (i.e. ratios 11:1 and greater) stimulus-pairs showed that these variances were heterogeneous. The variability was in fact significantly ($P < .01$) less for judgments involving large stimulus-differences than for judgments of relatively small differences.

It is not possible to state, on the basis of the present study, whether these scales have ratio properties. The validity of psychological scaling techniques is often assumed on the basis of the operations by which numbers are assigned to events, and is evaluated in terms of the internal consistency of the judgments. With respect to these criteria, ratio properties could be claimed for these scales, but perhaps would be of doubtful value in view of the demonstrated context effects. Garner has recently suggested that this is an extreme viewpoint since it argues for ratio properties of scales only on the basis of the stated operations and the reliability and internal consistency of the results.[8] Guilford and Dingman's comparison of the constant-sum judgments with fractionation-judgments, on the other hand, lend some evidence to the view that both reliable and valid judgments may be obtained with this method.[9] It will be necessary to carry out further studies to separate questions of the fundamental soundness of the method from the problem concerning errors which crop up under specific conditions.[10]

---

[7] A question raised by Guilford and Dingman, op. cit., 409.
[8] W. R. Garner, Context effects and the validity of loudness scales. J. exp., Psychol., 48, 1954, 218-224.
[9] Guilford and Dingman, op. cit., 398-404.
[10] See R. S. Woodworth and Harold Schlosberg, Experimental Psychology, 1954, 238-246.

MOTION PERCEIVED WHILE VIEWING ROTATING
STIMULUS-OBJECTS

By THOMAS MULHOLLAND, Clark University

During experiments with a rotating stimulus-object, many different motions were reported by Ss viewing the object against a homogeneously illuminated background.[1] Because of the theoretical interest of these phenomena, the initial observations were repeated under different conditions in the hope that some of the determinants of the perceived motions could be grossly specified. Four variables were studied: time of exposure to the rotating stimulus-object, object-background brightness-gradient, shape of the stimulus-object, and speed of rotation.

Apparatus. The apparatus had three essential parts: a stimulus-object, a homogeneously illuminated background, and a device to rotate the object in a horizontal plane at a constant speed. The stimulus-objects, made of stiff cardboard, were approximately 3.5 in. in diameter, and were usually attached at the axis of vertical symmetry to a black shaft 0.25 in. in diameter. The shaft was attached to a power-source which rotated the objects at various constant speeds. The background (8×11 in.) was placed 8 in. behind the object. The brightness of the background could be varied from very light to very dark. The visual display, consisting of rotating object and background, was enclosed in a black box which was lined with black velvet. In the front of this box was a rectangular opening 5×3 in. The S, sitting 9 ft. in front of the display in a darkroom, saw only the rotating object and the background framed by the rectangular opening.

Effect of time of exposure. The object was a 2-in. square rotated at 78 r.p.m. about its vertical axis. S observed the rotating object and described the kinds of motion he was experiencing. The reports were given, in response to a signal, at 10-sec. intervals for a period of 300 sec. The measures taken were the frequency of occurrence of reports of each kind of motion, and the time of first report of each motion. Ten Ss were tested.

The phenomena described by the Ss may be grouped into two classes, each with two sub-classes. There was three-dimensional motion of two kinds: *rotation*—the object is seen to rotate around its vertical axis, the perceived motion being clockwise (*CW*) or counterclockwise (*CCW*); and *oscillation*—no complete rotations are seen; the object seems to make a half turn *CW*, then a half turn *CCW*, then back again, alternately left to right, right to left, and so forth.

There was also two-dimensional motion of two kinds: *expansion-contraction*—no rotary or oscillatory motions are seen; the object seems to change its shape, contracting and expanding; and *reversal of figure and background*—this motion is simi-

lar to expansion-contraction, but the identity of the object is lost; the physical background is seen as a curtain-like figure which opens and closes over a seemingly partially visible ground, which is really the rotating physical object.

The various kinds of motion appeared in a definite sequence: rotation, oscillation, expansion-contraction, and reversal of figure-ground. The differences between the perceived motions with respect to the time they were first reported were tested by a nonparametric method and found to be significant beyond the 1-% level of confidence. The perceived motions also had significantly different frequencies of occurrence (beyond the 1-% level of confidence) as tested by the same method. The order, from highest frequency to lowest, was: rotation, oscillation, reversal of figure and background, and expansion-contraction.

Effect of object-background. In a series of experiments concerning the effect of auditory stimulation on visual brightness, a rotating object was employed as part of the apparatus. This object generated a unique perceived motion. The observations reported here were made under the control conditions, with no experimentally controlled auditory stimulation. The rotating object was *T*-shaped. At one end of the crossbar of the *T* was a disk which was white on both sides; the disk at the other end of the crossbar was black on both sides. The crossbar itself was a medium gray in the first study reported and alternate black and white horizontal stripes in the

TABLE I

MEAN FREQUENCY OF OCCURRENCE OF WHITE-FRONT AND BLACK-FRONT OSCILLATION
AT EACH OF THREE BACKGROUND-BRIGHTNESSES

Perceived motion	Background-brightness		
	Light	Medium	Dark
white-front	0.5	3.4	9.6
black-front	7.2	2.1	2.5

second study reported. The disks were 0.5 in. in diameter; the crossbar was 0.25 in. wide and 2 in. long. The object was attached at the middle of the crossbar to a black vertical shaft and rotated at 78 r.p.m. The brightness of the background was varied by changing its reflectance. The incident illumination on the background was between 0.05 and 0.10 foot-candles. Three independent groups of 12 Ss each were tested. Each group was tested with a different gray background. Each S reported on his perceived motions every 10 sec. for a period of 300 sec. The measures taken were the frequency of occurrence of reports of the different kinds of motion. Of major interest here is the oscillating motion. With the object used in this study, the oscillation may be perceived in two mutually exclusive ways. Sometimes the *white disk* seems to swing back and forth *in front of* the vertical axis of rotation, while the black disk seems to swing behind the axis of rotation; this motion is called *white-front*. At other times, the *black disk* seems to swing back and forth *in front of* the vertical axis of rotation, while the white disk seems to swing behind the axis of rotation; this motion is called *black-front*. As the background varies from light to dark gray, the number of reports of white-front increases, while the number of reports of black-front decreases. Table I presents the results for the three groups of Ss tested at different background-brightnesses. The data were evaluated in two analyses of variance, one for white-front, the other for black-front. The differences be-

tween background-brightnesses with respect to the number of reports of the oscilla-
tion motion were significant beyond the 1-% level for both kinds of oscillation.

A second study of this perceived motion also was carried out with the same ap-
paratus and procedure. Four Ss were tested with each of 9 backgrounds ranging in
unequal steps from light to dark gray. The results are presented in Table II. As the
background becomes darker, the number of reports of white-front increases, while the
number of reports of black-front decreases. The differences among backgrounds with
respect to the number of reports of each kind of oscillation were tested nonpara-
metrically. The differences were significant beyond the 1-% level for white-front,
and beyond the 5-% level for black-front.

Effect of speed of rotation. The stimulus-object was the Ames trapezoidal 'window'
and the rectangular 'window' proportionally reduced from the one reported in Ames'

TABLE II

MEAN FREQUENCY OF OCCURRENCE OF WHITE-FRONT AND BLACK-FRONT OSCILLATION
AT EACH OF NINE BACKGROUND BRIGHTNESSES

Perceived motion	Approximate background reflectance							
	84	64	47	30	18	17	11	4
white-front	0.0	0.5	2.0	4.5	6.0	7.5	16.8	15.5
black-front	14.3	7.5	1.0	1.3	3.3	0.3	0.0	0.3

monograph.[2] The figures were made of stiff cardboard and were 3.5 in. along the
longest diagonal. The speeds of rotation were 12–15 r.p.m.; 27–30 r.p.m.; 90 r.p.m.
The observations were made by the author and a colleague. Both monocular and
binocular regard were employed.

At 12–15 r.p.m., the perceived motions were predominantly those reported by
Ames. The rectangular stimulus-object was seen as rotating and the trapezoidal
stimulus-object was seen as oscillating. Other motions also were perceived. The
rectangle was seen oscillating in unison with the trapezoid. At other times, the rec-
tangle was seen oscillating contrary to the trapezoid. When the rectangle seemed to
turn *CW*, the trapezoid seemed to turn *CCW*, and so forth. The rectangle also was
seen to expand and contract in the frontal-parallel plane. At 27–30 r.p.m., the rec-
tangle was seen as rotating and the trapezoid as oscillating but these motions were
not as frequent nor as persistent as they were at a slower speed. At other times the
rectangle was seen to oscillate and the trapezoid to rotate. Expansion-contraction of
both figures in the frontal-parallel plane also was perceived. At 90 r.p.m., the mo-
tions described by Ames, *i.e.* rotating rectangle and oscillating trapezoid, were seldom
perceived. Perceived oscillation of both stimulus-objects was very frequent and per-
sistent. Expansion-contraction of both stimulus-objects also was seen. In sum, it is
clear that the frequency of occurrence of the perceived motion is a function of
the speed of rotation of the stimulus-object.

Effect of shape. The rotating objects were solid-plane cutouts ranging in size
from 2–5 in. in diameter. The speed of rotation was 78 r.p.m. Observations were
made by the author.

[2] Adelbert Ames, Visual perception and the rotating trapezoidal window, *Psychol.
Monogr.,* 65, 1951, (No. 7), 3-5.

For a square rotated about the vertical axis of symmetry, S perceived rotation CW and CCW, oscillation, expansion-contraction, reversal of figure and background. For a trapezoid rotated about the midpoint vertical, rotation CW and CCW, oscillation, expansion-contraction, and reversal of figure and background were observed. Oscillation was the predominant motion. A square rotated about the axis formed by a vertical side gave rotation, oscillation, and expansion-contraction. A trapezoid rotated about the axis formed by the shorter vertical side gave a very persistent oscillation, as well as rotation and expansion-contraction. A circular disk rotated about the vertical diameter gave frequent and persistent expansion-contraction and reversal of figure and background, as well as rotation and oscillation. The circular disk rotated about a vertical tangent gave oscillation (very frequent), rotation, and expansion-contraction. An ellipse rotated about the vertical axis of symmetry gave expansion-contraction and reversal of figure and background (very frequent), rotation, and oscillation. In general, it may be noted that oscillation is predominant with asymmetrical objects such as a trapezoid, or with an object which rotates about a vertical axis other than the axis of symmetry.

Summary. Four studies were conducted in order to specify some of the determinants of various kinds of motion experienced while viewing rotating stimulus-objects. The results indicate that the determinants of the motion are complex configurational factors, such as the brightness-gradient of the object-background, and shape. Speed of rotation of the object and time of exposure to the rotating object also play a rôle. These findings point up the difficulty of interpreting such perceived motions on the basis of a simple hypothesis of past-experience which does not take such configurational factors into account.

LEARNED INHIBITION OF 'SOUND-INDUCED' SEIZURES IN THE RAT

By A. E. HARRIMAN, Franklin and Marshall College, and
A. M. BRIAN, JR., Louisiana State University

Maier has proposed that the frustration which is rapidly generated in conditions of intense, inescapable, auditory stimulation may precipitate convulsive seizures in the rat.[1] Others have rejected the inference that frustration provokes seizures, holding that the behavior is comprised of reflexes which are excited by severe stimulation of the auditory centers.[2] The question of whether sound acts indirectly or directly in the production of convulsive seizures is examined in the present study. The hypothesis is tested that rats which learn that an opportunity for escape will be permitted at some point during exposure to intense sound will develop a tolerance for progressively lengthened periods of sound-stimulation, while rats subjected to the same progressively lengthened periods of sound but given no opportunity for escape will fail to develop a comparable tolerance.

Apparatus. A plywood box (20 in. long, 10 in. wide, and 10 in. high), painted flat black and partitioned into two identical chambers, was used. The front chamber was equipped with a glass-paneled, hinged door and with a 10-v. bell. The partition accommodated an externally controlled sliding door which permitted escape from the front to the rear chamber.

Subjects. One male and 23 female rats of the Wistar strain, 240-250 days of age at the beginning of the experiment, met a criterion of susceptibility to seizures. These animals manifested a minimum of three seizures, involving both the tonic and clonic phases, in five 120-sec. trials. On each trial, separated from the next by a 24-hr. interval, a rat was placed in the front chamber of the apparatus and subjected to the sound of a 10-v. bell.

The Ss were divided randomly into two groups—an experimental group of 11 rats, and a control group of 13 rats. The mean latencies (time between onset of the sound and onset of the seizures) on the criterion trials were 66.7 sec. for the experimental group and 63.1 sec. for the control group. The difference between these means was insignificant ($t = 1.74$, 22 df). The rats in both groups lived in individual cages and had free access to food (Purina Dog Chow) and water.

Procedure. Two trials, separated by an 8-hr. interval, were given daily to each rat. Convulsive responses elicited from the rats during stimulation fell into two categories. In the first category were listed those episodes characterized either by undirected running or by clonic attacks. In the second category were included only responses involving both the clonic and tonic phases of the full seizure pattern.

Following a period of exploration of the plywood box, each S in the experimental group underwent an initial training phase of 15 trials. On each trial, the rat was individually placed in the front chamber and subjected to the sound of the bell for

---

a maximum of 20 sec. After 15 sec. of stimulation, however, the sliding door between the chambers was raised, and the animal was given 5 sec. in which to leave the front chamber before the cessation of sound. If the animal entered the rear chamber during the 5-sec. interval, the sound ceased immediately, and the door was lowered. Those rats which passed into the rear chamber were allowed to remain there for 3 min. before being returned to their cages. The others were removed from the front chamber at the end of the 20-sec. period. By the 15th trial, all of the experimental animals entered the rear chamber before the 20th second.

Beginning with the 16th trial, the time of auditory stimulation prior to the elevation of the door was increased by 2 sec. after every second trial. Otherwise, the conditions were the same as before. From the 37th to the 60th trials, the time prior to opportunity for escape was increased by 3 sec. after every second trial. From the 61st to the 80th trials, the time was increased by 5 sec. after every pair of trials. Hence, on the final pair of trials, each experimental rat received 121 sec. of exposure to sound before the door was raised.

For members of the control group, there was no opportunity for escape at any time during the experiment. Placed individually in the front chamber, each control S was stimulated for an appropriate period and then removed from the chamber. During the initial training phase for the experimental animals, the controls received fifteen 20-sec. periods of stimulation. Following this phase, as the length of the exposure-periods was increased for the experimental rats, the periods were lengthened from the base of 20 sec. by identical intervals for the control animals. On the 77th and 78th trials, the controls received 121 sec. of exposure.

Results. Both groups received approximately the same amount of sound-stimulation. The mean total exposure for the experimental group was 4167 sec. in 80 trials, and for the control group it was 4154 sec. in 78 trials. The frequencies of attacks characterized by undirected running or by clonic seizure were, however, by no means equivalent for the two groups. The control group exhibited a mean total of 26.8 such responses during their 78 trials; the mean total for the experimental animals was only 7 such responses during 80 trials. The difference was significant at a level of confidence beyond the 1% level ($t = 4.85$ with 22 df). A further test showed no reason to reject the null hypothesis with regard to homogeneity of variance.

A significant difference between the groups also was obtained for the occurrences of full clonic-tonic seizures. The control group manifested a mean total of 19.5 clonic-tonic seizures, and all the animals in the group contributed to the total (the smallest number of seizures in any animal being 5 and the largest being 37). The rats in the experimental group exhibited a mean total of 1.5 seizures, and only six of the animals were so affected (one animal having 10 seizures, another 2 seizures, and four others showing only 1 seizure each). The difference between the means for this category of response was significant beyond the 1% level ($t = 5.00$, with 22 df), but the variances were heterogeneous. An 'adjusted t-value' therefore was derived (adjusted $t = 3.06$, with 22 df) by the method of Cochran and Cox, which indicated that the difference was significant beyond the 1% level.[3]

A sharp rise in the frequency of seizures in the control rats following the 32nd trial and continuing throughout the remainder of the experiment may be seen in Fig. 1. For the experimental rats, the appearance of a significant tendency toward seizures was delayed for another 24 trials, and the increase in frequency of seizures which then occurred was not sustained. While the control group had a mean of 4.7

---

[3] W. G. Cochran and G. M. Cox, *Experimental Design*, 1950, 92-93.

seizures during the final block of trials, the mean frequency of seizures for the experimental rats decreased to 0.2.

Discussion. Although the criterion trials demonstrated that the two groups were equally susceptible to seizures, the groups differed significantly during the experiment
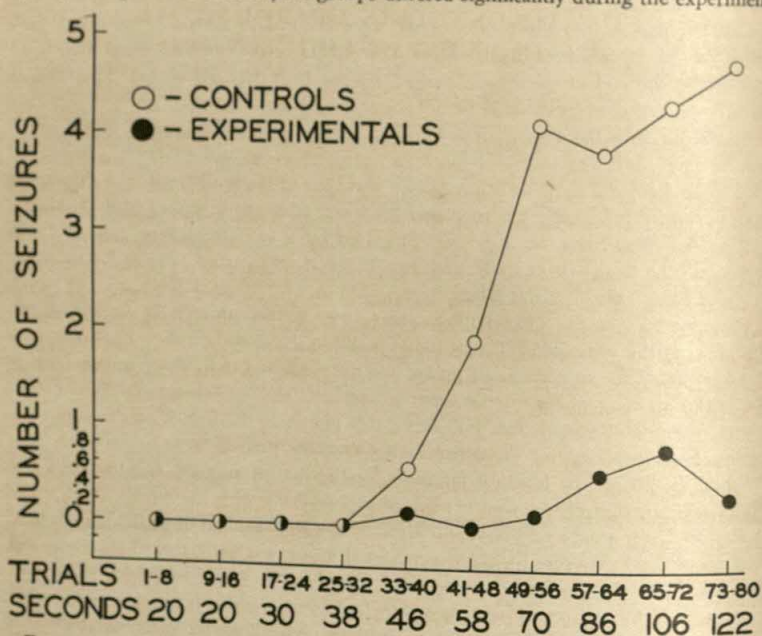
FIG. 1. NUMBER OF SEIZURES IN THE CONTROL AND EXPERIMENTAL GROUPS

in the manifestation of seizures. This difference presumably occurred because the consistent presentation of an opportunity for escape at some point on each trial enabled the experimental rats to develop a tolerance for intense auditory stimulation. By the end of the study, this tolerance was of sufficient strength to permit the experimental rats to withstand periods of sound-stimulation which were approximately twice as long as those sufficient to produce seizures during the criterion trials. These results lend support to Maier's view that sound is an indirect rather than a direct cause of the seizures.

Summary. This study reports a test of Maier's premise that sound generates frustration which evokes seizures in the rat and, hence, acts as an indirect rather than a direct cause of this behavior. Two groups of rats which exhibited considerable susceptibility to seizures were utilized. Following progressively lengthened periods of intense sound, members of an experimental group were given opportunity for escape. Members of a control group, subjected to equivalently lengthened periods, were given no such opportunity. Although both groups received approximately equal periods of auditory stimulation throughout the study, the experimental group manifested significantly fewer partial and full seizures than did the control group. This finding is interpreted as support for Maier's view.

THE RELATIVE EFFICACY OF VARIOUS TYPES OF STIMULUS-OBJECTS IN DISCRIMINATIVE LEARNING BY CHILDREN

By Allen D. Calvin and L. Thomas Clifford, Michigan State College

A problem of growing importance in the field of discriminative learning is the relative effect of various types of stimulus-objects. While a substantial amount of research has been done in this field with subhuman Primates,[1] comparatively little has been done with children.[2] A thorough exploration of the effect of different stimulus-characteristics on discriminative learning in children should not only add to our knowledge of the variables affecting children's learning but it should also enable us to make phylogenetic comparisons that will be of aid in our interpretation of animal learning. The present study was undertaken, therefore, as a step toward acquiring that knowledge.

METHOD AND PROCEDURE

Apparatus. An exposure-apparatus, constructed of white cardboard, was used to present the materials. Its central panel, 22 in. wide and 28 in. high, which could be raised and lowered by *E*, was the exposure-screen. The stimulus-cards were placed two at a time behind the closed screen and in front of identical red cups under one

[1] H. F. Harlow, Studies in discrimination learning by monkeys: VI. Discrimination between stimuli differing in both color and form, only in color, and only in form, *J. gen. Psychol.*, 33, 1945, 225-235; D. R. Meyer, and H. F. Harlow, The development of transfer of response to patterning by monkeys, *J. comp. & physiol. Psychol.*, 42, 1949, 454-462; J. M. Warren, Additivity of cues in visual pattern discrimination by monkeys, *ibid.*, 46, 1953, 484-486; J. M. Warren, and H. F. Harlow, Learned discrimination performance by monkeys after prolonged postoperative recovery from large cortical lesions, *ibid.*, 45, 1952, 119-126; K. L. Chow, Stimulus-characteristics and rate of learning visual discriminations by experimentally naïve monkeys, this JOURNAL, 66, 1953, 278-282; J. Cole, The relative importance of color and form in discrimination learning in monkeys, *J. comp. & physiol. Psychol.*, 46, 1953, 16-18; H. W. Nissen, and W. O. Jenkins, Reduction and rivalry of cues in the discriminative behavior in chimpanzees, *ibid.*, 35, 1943, 85-95; H. W. Nissen, and T. L. McCulloch, Equated and non-equated stimulus situations in discriminative learning by chimpanzees; III. Prepotency of response to oddity through training, *ibid.*, 23, 1937, 377-381.

[2] There have been a number of studies utilizing various matching techniques for determining children's interest in color, form, etc. While not directly related to learning problems they do provide relevant background material which can be of considerable aid in selecting particular stimulus-dimensions for specific age-groups. An excellent review of these studies appears in Chapter 12 of M. D. Vernon *Visual Perception*, 1937, 178-189.

of which the reward (a plastic toy) was concealed. The cards and cups behind them were placed 4 in. apart and even in the frontal plane. S sat at a table with the apparatus before him. E operated the screen and set the stimulus-cards, cups, and rewards in their proper positions from behind it.

Subjects. Forty first grade students from the public school of East Lansing were used as the Ss. They were divided randomly into five groups of eight each: Group I had to discriminate between a plain black and a plain white stimulus-card; Group II, plain white vs. a white card with a small silver star in the center; Group III, a black card with a small silver star in the center vs. a white card with a small silver star in the center; Group IV, dark green vs. light green; and Group V, blue vs. dark green.[3] A balanced experimental design was employed with all groups.

Each S was randomly selected from the class by the teacher to solve one of the problems previously determined by E. The teacher knew nothing of the nature of the problem. While walking to the testing room, S was asked such questions as, "Do you know where we are going?" and "Do you know what we are going to do?" as a check on possible intercommunication between Ss about the experiment. It was found that some of the Ss had talked to each other about the game but not about the actual problem involved.

Upon entering the testing room, S was seated at the exposure-screen and given the following instructions.

We are going to play a little game. Behind this screen are two cups, and under one of them there is a toy. If you can guess which cup the toy is under, you can keep it. We want this game to be fun for everybody; so, we would like you to help us by keeping it a secret.

E then placed the stimulus-cards, cups, and the reward in the previously determined positions and then at a signal for S's attention, raised the screen. S chose one of the cups. If the correct cup was chosen S was immediately handed the reward. If an incorrect choice was made, the cup was raised to reveal that there was nothing under it, and the positive cup was raised to show that it did have a toy under it. The screen was then lowered, and the reward and stimulus-cards replaced according to a predetermined design in which a given card did not appear on the same side more than twice in succession.

The criterion of learning was 10 consecutive correct trials. The test was terminated at the first error after the thirtieth trial. Ss not learning in 30 trials were assigned arbitrarily a score of 40.

RESULTS AND DISCUSSION

Table I presents the learning scores for the various groups. The performance of Group V is markedly inferior to the other groups. Because of the difficulty of meeting the assumptions of a parametric test, the Mann Whitney non-parametric U-test was used.[4] The difference between Group V and Group II was significant beyond the 5-% level. The differences between Group V and Groups I, III, and IV

---

[3] Our light green approximates Nile Green No. 60; our dark green approximates Shamrock Green No. 64; and our blue approximates Italian Blue No. 82 according to the color chart presented in the unabridged second edition of Webster's *New International Dictionary*, 1949, 528.

[4] H. B. Mann and D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, *Ann. Math. Stat.*, 18, 1947, 50-60.

fell just short of significance. The differences between Group V and the rest of the groups combined was significant beyond the 5-% level. There were no other significant differences.

The results indicate that it is more difficult for children to discriminate between different hues than between achromatic stimuli, different brightness levels of the same hue, or patterns. On the basis of the Ss' reports, the possibility of ascribing the poorer performance of Group V to the differences in hue being subliminal can

TABLE I

MEDIAN NUMBER OF TRIALS REQUIRED BY THE VARIOUS GROUPS OF Ss

	Group				
	I	II	III	IV	V
Trials	16	15	17	16	40

definitely be ruled out. All Ss were questioned at the conclusion of their experimental session, and all of them could identify which stimulus-card was blue, which was green, etc. Their poor performance was due to an inability to realize that the stimulus-cards were cues for the solution of the problem. In the authors' opinion, this is due to the fact that although both cards were distinctly different in hue, the Ss still conceptualized them as *colored cards* rather than as a blue card or a green card. This generalized concept *colored cards* thus served to prevent the Ss in Group V from using the differential hues as cues.[5]

These findings are not in accord with the results usually obtained by experimenters working with subhuman Primates. Since the present experimenters were primarily interested in studying these processes in children rather than in making phylogenetic comparisons, the differences between our findings and theirs could possibly be due to differences in experimental procedure rather than real phylogenetic differences. A study now underway which attempts to duplicate the experimental conditions of previous subhuman Primate experiments more closely seem to indicate, however,

---

[5] M. G. Colby and J. B. Robertson (Genetic studies in abstraction, *J. comp. & physiol. Psychol.*, 33, 1942, 385-401) believe that color is perceptually analogous to genus while form is analogous to species; however, they point out that Brian and Goodenough hold that the converse is true. G. R. Brian and F. L. Goodenough (The relative potency of color and form perception at various ages. *J. exp. Psychol.*, 12, 1929, 197-213) state on p. 213: "When the child first begins to respond to partial elements in a situation he tends to select those which are common to many features of his environment. Objects are grouped by genus rather than by species. Thus all time pieces become 'clocks' or 'tick-tocks'; robins, crows, ducks, and hens are alike 'birds' or 'chickens,' and so on. In making these early classifications, structure and function are of major importance, while color can in most instances be ignored. A little later on, as facility in making gross classifications increases, the child becomes preoccupied with those attributes of a given object or situation which make for the differentiation of species within a given genus. Under these circumstances, color discrimination becomes more important. Still later, when a level of development has been reached at which the subject is able to shift his attention at will from one aspect of a situation to another, a new process of organization and evaluation of these partial elements takes place." This explanation is somewhat akin to ours; however, our results indicate that color *per se* can also be used generically which tends to support the position of Colby and Robertson mentioned previously.

that this is a real difference and not due primarily to variation in experimental procedure. It appears most likely that the difference lies in the conceptual approach to the problem as mentioned previously.

SUMMARY AND CONCLUSIONS

The effect of different stimulus-characteristics on discriminative learning in children was studied. It was found that it was significantly more difficult to discriminate between different hues than to discriminate between achromatic stimuli, different brightness levels of the same hue, or pattern stimuli. These findings are discussed and a possible explanation for the relative difficulty of the hue problem is presented.

INTERSENSORY TRANSFER IN THE DISCRIMINATION OF FORM

By Henry F. Gaydos, Quartermaster Research
and Development Center

Form-discrimination in the visual or the tactual-kinesthetic modalities has been studied extensively, but there has been no systematic investigation of the transfer of form-discrimination from one modality to the other. Ryan has reviewed a number of studies dealing with intersensory relationships, but they are concerned mainly with changes in the sensitivity of one sensory system during stimulation of another.[1] Worchel has studied form-perception in blind and in blindfolded Ss.[2] His results show that blind Ss can tactually discriminate and recognize form as well as seeing Ss, although drawings by blind Ss of tactually perceived forms correspond very poorly with visual perceptions of the same forms. From this result it appears that the similarity between visual and tactual form, and hence transfer, depends upon the development of both modalities. A recent study by Semmes, Weinstein, Ghent, and Teuber, comparing the performance of brain-injured Ss with that of normal Ss on tactual tasks, indicated some positive transfer between touch and vision.[3] This finding, however, was quite incidental to the main purpose of the investigation. The present experiment was designed to study the transfer of form-discrimination from each modality to the other.

Subjects. The Ss (86 in number—40 women and 46 men) were students (graduates and undergraduates) and faculty members.

Stimulus-objects. Twelve stimulus-objects, cut from $1/8$-in. Masonite, and the names assigned to them are shown in Fig. 1. These names were chosen to identify the forms because they are familiar to most people, and presumably easy to remember.[4] The size of each piece was approximately 2 in. across in its largest dimension. The notch appearing at the bottom of each piece served as a reference-point by means of which the spatial orientation of the shape was made known to S. Each stimulus-object always was presented with the notched side toward him. Since

[1] T. A. Ryan, Interrelations of the sensory systems in perception, *Psychol. Bull.*, 37, 1940, 659-698.

[2] Philip Worchel, Space perception and orientation in the blind, *Psychol. Monogr.*, 65, 1951, (No. 332), 1-28.

[3] Josephine Semmes, Sidney Weinstein, Lila Ghent, and H. L. Teuber, Performance on complex tasks after brain injury in man: analysis by locus of lesion, this JOURNAL, 67, 1954, 220-240.

[4] Part of the initial learning-effort was spent merely in memorizing these names. It is quite possible, therefore, that some errors in the early trials resulted from a failure to remember names rather than from a confusion of shapes. This problem might have been avoided by using numerals or letters of the alphabet to label the shapes. It was feared, however, that the tendency to think of individual numbers or letters as having well established positions in a series might also retard learning if the objects to be designated by numbers or letters were presented in random order. For this reason, one-syllable names of men were chosen for convenient labels.

Masonite has one smooth side and one rough side, the possibility of confusing S with an inverted shape was avoided.

Procedure. The Ss learned to recognize each shape through one sense-modality, and then attempted to identify the same shapes through a different modality. The initial learning was carried to a criterion of two consecutive errorless trials, and the test for transfer was carried to one errorless trial. Group I learned first by touch, and was then tested for ability to recognize the shapes by sight. Group II learned by sight first, and was tested for transfer to recognition by touch. Under these con-
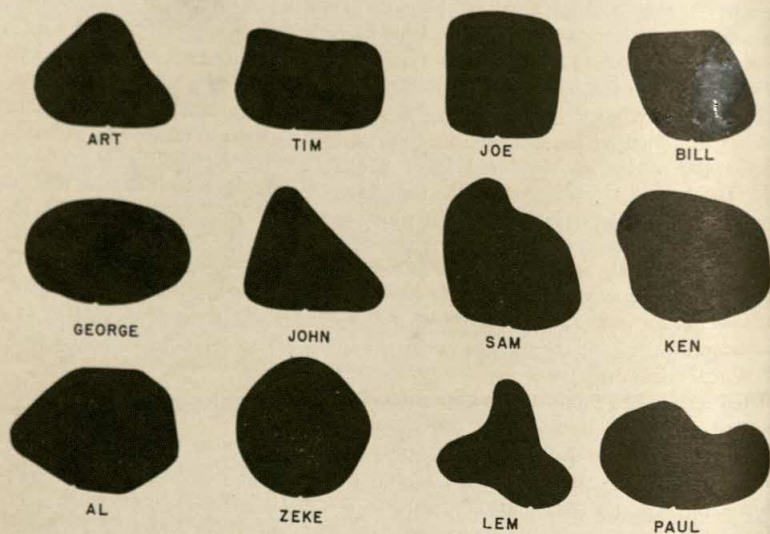
FIG. 1. THE STIMULUS-OBJECTS

ditions, each group served as the control for the other. The Ss were divided equally between Groups I and II in such a way as to equate the groups as nearly as possible for age, sex, and educational level.

During the experiment with Group I, each S was seated before a screen which hid the shapes from view while he manipulated them. The shapes were introduced by name, one at a time, and S was allowed to explore the contours with both hands. After the introductory trial, S was required to try to name each piece as it was presented, and the correct name was given by E after each response. The shapes were presented in a different order on each trial, and no sign was given to tell when one trial ended and the next began. The presentations continued until the criterion of learning was attained, and then the stimulus-objects were presented visually, by the same procedure described above, until one errorless trial was achieved. The procedure for Group II was the same as for Group I, except that initial learning was visual and the relearning tactual.[5]

---

[5] When time relationships are considered, the difficult question arises as to what constitutes equivalence of sensory contact in the two modalities. While a fraction

Results. Performance was recorded in terms of trials and errors. While the relationship between the two measures was not perfect, the correlations (0.81 for Group I and 0.96 for Group II) were close enough to warrant using either measure alone as an index of performance. Statistical tests based on both types of data gave similar results. The data are summarized in Table I. In initial learning, visual discrimination was slightly superior to tactual discrimination, although *t*-tests indicated that neither the difference in trials nor the difference in errors approached statistical significance. When the relearning scores for each group are compared with

TABLE I

LEARNING AND RELEARNING IN EACH MODALITY

Modality	Trials			Errors		
	learning	relearning	saving (%)	learning	relearning	saving (%)
Touch	7.93	2.05	74	39.65	4.58	88
Vision	6.69	1.05	84	34.93	2.02	95

the initial learning scores of the other, it is obvious that there was marked transfer from each modality to the other. In each modality the difference between learning and relearning for both trials and errors was significant beyond the 1-% level of confidence.[6]

When relearning scores are compared, it appears that visual recognition after initial learning by touch required fewer trials and fewer errors than did tactual recognition after visual learning. The differences are significant at the 1-% level of confidence and indicate that transfer-effects were greater from touch to vision than from vision to touch. This result is not unexpected, since learning to identify form by touch alone, a relatively unfamiliar process, may have prompted the *S*s of Group I to explore the shapes more thoroughly. The *S*s of Group II may have been able to learn initially on the basis of more generalized properties which may not always have been sufficient to allow recognition by touch.

Correlations between learning and relearning are quite low. All are lower than 0.30, and none is reliably different from zero. Accordingly, the efficiency of transfer seems to depend not on the rate, but rather on the degree, of initial learning. A possible interpretation of this result might be that whatever transfers across modalities is in the nature of a group of specific concepts of the different shapes, and not merely a general practice effect. In the latter event, higher correlations between the scores on the initial and subsequent tasks would be expected.

Certain observations of a more qualitative nature are worth noting. Verbalization seemed to play an important part in the learning for many of the *S*s. Several kinds of verbal association were employed in learning to differentiate and recognize the shapes. First, certain shapes were associated with common geometric figures, such as

of a second may suffice for the eye to perceive a given form in its entirety, manual exploration is generally a slower process. In view of this problem, no time-limits were set for either tactual or for visual contact with the stimulus-objects.

[6] Where the variances were not homogeneous, the test was applied using half the normal *df* as suggested by G. W. Snedecor, (*Statistical Methods*, 1946, 82-83).

square or triangle. When familiar geometric terms were not adequate to describe all the figures, Ss would resort to associations such as, "Lem is like a floppy hat." These memory-aids varied in complexity, but seemed to be associated only with perceived shape. A more complex type of association was sometimes employed to link the shape and name together. For example, "John is a triangle because 'John' reminds me of 'John's other wife,' which reminds me of the 'eternal triangle'." The Ss who were able to use this kind of association seemed to be faster learners.

Some differences in apparent size were reported in connection with vision versus touch. The objects generally felt bigger than they looked, and contour features seemed more prominent to touch than to vision. This finding may have some bearing on the differences in degree of transfer discussed earlier. If the bumps and hollows were more prominent to the touch, then perhaps those who learned by this modality were able to perceive more of the salient features of each stimulus-object than those who learned by sight.

THE EFFECT OF SLEEP-DEPRIVATION UPON THE THRESHOLDS OF TASTE

By ERNEST FURCHTGOTT and WARREN W. WILLINGHAM,
University of Tennessee

Though a number of studies have been made of the effects of sleep-deprivation on various sensory functions,[1] none has dealt with the effect on taste. The present study attempts to fill this gap.

METHOD AND PROCEDURE

Subjects. The Ss in this study (18 men) were college students ranging between the ages 21–30 yr., who were paid for their services.

Procedure. Absolute thresholds were determined for sour (HCl), salt (NaCl), and sweet (sucrose) by a modification of the method of limits.[2] S was presented with two beakers, one containing distilled water and the other a subliminal concentration of the test-solution. S compared the two by sipping, going back and forth between them. The concentration of the test-solution was increased in successive comparisons until S could detect a difference at a given concentration and could identify the beaker with the test-solution at three successive trials. The beakers containing the water and the test-solution were randomly alternated left and right in presenting them to S. Three ascending series, one with each of the three test-solutions, were conducted in randomized order during an experimental session. Five sessions were held with every S.

The average reliabilities of the just noticeables (*RL*s) of all the Ss for the five sessions were 0.95, 0.94, and 0.86 for sour, salt, and sweet, respectively;[3] an indication that our method was highly reliable.

The normal thresholds were obtained a day before the experimental periods began and again immediately preceding the period of sleep-deprivation which ran to 72 hr. Thresholds were obtained from the Ss at the 24-, 48-, and 72-hr. periods.

RESULTS

Table I gives the average thresholds of our Ss for each of the taste qualities used (sour, salt, and sweet) at every experimental period. The data for each quality were

[1] Nathaniel Kleitman, *Sleep and Wakefulness,* 1939, 283-321. A. S. Edwards, Effects of loss of 100 hours of sleep, this JOURNAL, 54, 1941, 80-91; Victor Goodhill and D. B. Tyler, Experimental insomnia and auditory acuity, *Arch. Otolaryng.,* 46, 1947, 221-224.

[2] C. P. Richter and Alice MacLean, Salt taste thresholds of humans, *Amer. J. Physiol.,* 126, 1939, 1-6.

[3] R. L. Ebel, Estimation of the reliability of ratings, *Psychometrika,* 16, 1951, 407-424.

subjected to an analysis of variance. The F-value for sour was 4.04 which is almost at the 1-% level ($P\ 0.01 = 4.19$). As the table shows, the increase in the threshold for sour (HCl) occurred the 24- and 48-hr. periods of deprivation. The F-value for salt (NaCl) was 0.51 and for sweet (sucrose), 1.36, neither being statistically significant.

Our results for salt and sweet are consistent with those of other investigators who showed that deprivation of sleep does not affect any of the sensory functions except pain—which it depressed.[4] Why sensitivity to sour showed a significant deterioration, we do not know. Since sensitivity to sour is much higher than to salt and sweet,

TABLE I
AVERAGE RLs AS A FUNCTION OF SLEEP-DEPRIVATION
(Concentrations expressed as grams of solute dissolved in 100 ml. of water)

Hours of deprivation	Sour (HCl)	Salt (NaCl)	Sweet (sucrose)
0	.00122		
24	.00123	.041	.30
48	.00162	.039	.33
72	.00164	.039	.37
		.036	.39

it may be assumed that general fatigue, and its associated decrease in attention, which is usually experienced in sleep-deprivation, will have a more pronounced effect on sour, the most acute taste, than upon the other qualities. A more speculative hypothesis is the association of a lower sensitivity to sour with the greater alkalinity of the body fluids that is reported during sleep-deprivation.[5]

---

[4] Kleitman, *op. cit.*, 306.
[5] *Idem*, 310.

ANXIETY AND DISCRIMINATIVE LEARNING

By HAROLD W. STEVENSON and IRA ISCOE, University of Texas

In a recent experiment in which college students were required to learn to discriminate one of three objects differing in size, great individual differences in speed of learning were found.[1] Highly variable performance is typical of most learning experiments with human adults, but it is of interest here that some Ss required up to 78 trials in learning a simple problem which the average S solved in 10 trials. Among the variables that might be related to the differences in performance is the level of anxiety with which Ss faced the task, for it has been found in a number of recent studies that anxiety-level is related to the ease with which Ss are able to make discriminations and to learn simple problems.[2] The present experiment was undertaken to investigate the effects of anxiety on the learning of a simple discriminative problem such as that described above.

Subjects. The Taylor Anxiety Scale was administered to 346 students in introductory psychology at the University of Texas, and on the basis of these scores high-anxiety and low-anxiety groups of 20 Ss each were formed.[3] The mean scores for the high and low groups were 29.4 ($SD = 4.4$) and 4.0 ($SD = 1.7$), respectively. The Ss were selected for the high-anxiety (HA) group if they had scores of 23 or above, and for the low-anxiety (LA) group if they had scores of 6 or below. There were 10 men and 10 women in each group.

Apparatus. The apparatus consisted of: (a) three 3-in. square blocks, each bearing a white cardboard square; (b) a tray for presenting the blocks; (c) a screen to shield E; and (d) a poker chip which served as the goal-object. The areas of the white squares were 3/4, 1, and 1 1/2 sq. in. On the tray were three wells in which the goal-object could be concealed. The tray and blocks were painted flat black.

Procedure. The experiment was performed several weeks after the Anxiety Scale was administered, and the Ss had little basis for connecting their selection with their prior test-performance. The HA and LA Ss were tested concurrently, and E did not know the classification of a particular S until after the trials were completed.

S was seated at a table opposite E who demonstrated how the blocks could be used to cover the holes on the board. The following instructions were then given: "I'm going to hide this poker chip and you are to find it as often and as consistently

---

[1] H. W. Stevenson and George Moushegian, Response-shift as a function of instructions and degree of training, this JOURNAL (in press).

[2] Examples of experiments in which differences were found are the following: E. R. Hilgard, L. V. Jones, and S. J. Kaplan, Conditioned discrimination as related to anxiety, *J. exp. Psychol.*, 42, 1951, 94-100; I. E. Farber and K. W. Spence, Complex learning and conditioning as a function of anxiety, *J. exp. Psychol.*, 45, 1953, 120-126; J. D. Lucas, The interactive effects of anxiety, failure, and intra-serial duplication, this JOURNAL, 65, 1952, 59-67.

[3] J. A. Taylor, A personality scale of manifest anxiety, *J. abnorm. soc. Psychol.*, 48, 1953, 285-291.

as you can. You will get one choice per trial." If S failed to understand the instruc-tions they were repeated, but no additional information was offered. The position of the blocks on the board followed the same random sequence for all Ss, and the poker chip was always placed under the block bearing the middle-sized square. The criterion for learning was five successive correct responses.

Results. With the 5 criterion trials omitted, the mean number of trials required to reach the learning criterion by the HA Ss was 28.0 trials ($SD = 37.5$) and by the LA Ss it was 5.8 trials ($SD = 7.0$). The respective medians were 10.0 and 4.0. The range of trials for the LA Ss was 0 to 28 and for the HA Ss it was 1 to 156. Because of the skewness of the distributions, the difference in learning scores between the groups was tested by Wilcoxon's nonparametric test for unpaired replicates.[4] The difference is significant at the 1-% level of confidence.

Discussion. These data reveal that the LA Ss were rather consistent in requiring few trials to learn, while the HA Ss required more trials and were quite variable. We may conclude that the level of anxiety is a significant variable in determining the ease with which a simple discrimination is learned, and also that it affects the variability in learning rate. The extreme differences in performance among the HA Ss poses an interesting problem in interpreting the rôle of anxiety in learning. Since there was no significant correlation within the HA group between anxiety-score and rate of learning ($P = 0.36$, $p > 10\%$), it must be assumed that the differences in performance within the HA group were due to the operation of variables other than anxiety. Although ratings of intelligence were not available, it is doubtful that the differences in learning rate could be attributable primarily to variations in intelli-gence. It seems more reasonable to assume, as Mandler and Sarason have done, that the differences in performance among the HA Ss were due to the varying ways in which particular Ss had learned to respond to their anxiety.[5] Mandler and Sarason have hypothesized that anxiety leads to poor performance because of the tendency for anxious Ss to make task-irrelevant responses. Thus, LA Ss tend to have less difficulty in learning than HA Ss, but HA Ss may, depending upon their past ex-perience in responding to anxiety with task-relevant responses, perform in an efficient manner.

[4] Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures,* American Cyanamid Co., Stamford, Conn., 1949, 1-16.

[5] George Mandler and S. B. Sarason, A study of anxiety and learning, *J. abnorm. soc. Psychol.,* 47, 1952, 166-173.

NOTES AND DISCUSSIONS

THE USES AND ABUSES OF 'CENTS' IN THE SCIENCE OF AUDITION

In psychology, the auditory stimulus is defined in terms of frequency or length of its sound waves. The sensation aroused by the stimulus is called tone, pitch, or note. A logarithm, *e.g.* the logarithm of 2, is a dimensionless number, whereas an auditory stimulus is always a cardinal number expressed in terms of physical dimensions—centimeters or seconds. The logarithm has no physical dimension and it is not a stimulus. Neither is a 'mill,' nor a 'cent,' nor a 'deci' of it a stimulus to a scientifically trained psychologist. The 'cent,' defined as 1/1200 log 2, was introduced by Ellis in his translation of Helmholtz's *Tonempfindungen* without Helmholtz's knowledge or permission; and the term never received Helmholtz's approval.

After Bach had published his "Twenty Four Pieces" under the title *Wohltemperiertes Klavier,* the manufacturers of pianos instructed their tuners to make each step from one key to the next as nearly as they could equal to the cycle ratio $1:2^{-12}$, which is approximately 1000:1059 (not a logarithm). This ratio, which is still being followed by piano tuners, is acceptable as no pianist expects any great exactness in tuning as long as the octaves are free from beats. The listener to the piano reacts to the other intervals (all mistuned—some more, some less) with the same magnanimity with which he tolerates in conversation the speech when pronounced by a speaker who has learned to speak the language with the nuances used in a different part of the world.

Assume now, that a psychologist becomes interested in measuring the amount of tolerance manifested in the hearing of orchestral music—a symphony. The natural requirement would be to determine beforehand by experiment what *percentage of cycles* is the *natural threshold* of deviation from the true intervals in those cases when the listener is given a chance himself to tune each separate interval without being forced by a manufacturer to accept what is offered. Such experiments were initiated by Stumpf.[1] The subjects were all staff members of the Royal Conservatory of Berlin, led by the famed violin virtuoso Joseph Joachim (1831-1907),

---

[1] Carl Stumpf and M. F. Meyer, Massbestimmungen über die Reinheit consonanter Intervalle, *Z. Psychol.* 18, 1898, 350 and 371. See also Meyer, The musician's arithmetic, *Univ. Mo. Studies,* 1929, 126-129.

and a psychologist, Carl Stumpf, who was also a highly accomplished musician although he did not make his living as such. The surprising result was that the measured threshold of deviation was *unsymmetrical* to the theoretically correct frequency and that the direction of this lack of symmetry appeared to be upwards for intervals larger than the so-called 'neutral third' (3.5 semitones), and downwards for intervals smaller than that. In other words, the Major Third, as tuned by the listener, was augmented by a few cycles and the Minor Third, diminished by a few cycles. The octave, contrary to the common idea of a want for having it exact, was surprisingly also preferred as augmented. When the Ss were asked if they had a guess justifying their revealed tendencies, they said that the smaller rising spans ought to express in addition to melody also 'dullness' and those larger rising spans ought to express also 'brightness' and that a little emphasis on these desiderata, dullness and brightness, is a good thing as long as it does not destroy nor seriously hamper the experience of 'melodiousness.'

Regrettably, no analogous experiments have ever been made with the same intervals in the *falling inflection,* when the upper tone is the constant one and the pitch of the lower one has to be chosen by the listener. What might be the choice then?

It took nearly 60 yr. before any psychologist undertook to repeat Stumpf's experiment. Ward attempted it in 1954.[2] He did not, however, expand the experiment in the direction of testing the falling inflection. He again held the lower tone constant and had his Ss choose the upper. Hence he could not answer our question. His work was even more limited than Stumpf's because he used no further interval than the octave.

On the other hand Ward had the advantage that he did not have to present to his Ss various tuning fork frequences from which they had to choose one. In Stumpf's time electronic audio-generators of which the frequency is changed by merely turning a knob did not exist and tuned forks were given to the Ss for choice. Ward's Ss were asked simply to set the generator until the higher frequency of the interval suited their fancy. The result was, however, the same as that obtained by Stumpf, *i.e.* the octave interval preferred was augmented by several cycles.

I am surprised, however, to find that Ward did not report the actual higher frequences chosen by his Ss. Why they were not reported is not explained and is hard to understand. He took the difference between the chosen frequency and the frequency of the true octave, but did not give this even in percentage of cycles for the reader to see. He computed the chosen

[2] W. D. Ward, Subjective musical pitch, *J. acoust. Soc. Amer.,* 26, 1954, 371b.

difference logarithmically into 'cents' and reported 'cents.' Was this done with the notion that a report of the differences in 'cents' would be *more scientific* than in cycle differences and their percentages of the higher frequencies for which I and the average reader would naturally look?

In Helmholtz's *Tonempfindungen* we find no mentioning of 'cents.' Ellis translated the book into English without consulting Helmholtz. Helmholtz recommended to the English reader Sedley Taylor's book, "Sound and Muisc" (1883), as "reproducing the essential contents" of his *Tonempfindungen*. Shortly before his death (1894), Helmholtz definitely stipulated that a new edition (which was published in 1895) should contain none of the additions that Ellis had incorporated in his translation. The reader can draw his conclusions from this fact.

Helmholtz, one of the founders of the *Zeitschrift für Psychologie*, was profoundly interested in sensory psychology. This fundamental attitude is so much overshadowed in the translation by Ellis that the naïve reader does not get the impression that Helmholtz was trying to write a psychological text. I wish, however, to restrict this discussion to the introduction of 'cents.' Of 'cents' Helmholtz says nothing, naturally, because 'cents' are not related to any psychological concept.

Ellis forces his unwanted 'cents' into the "Sensations of Tone" in several ways. I enumerate some of them: on p. 13 of the 1948 edition, Ellis rejects the idea of Helmholtz that a musical interval must *always* be expressed by the *integral cycle ratio* and substitutes "for some purposes by other numbers called cents." Helmholtz could never admit that! Cents are *irrational* numbers. On p. 446, Ellis confesses that, without the permission of Helmholtz, "cents have been . . . introduced into the text of this translation." On the same page, Ellis asserts that "it is necessary in order to have a proper conception of the [correct] interval . . . to determine the number of cents in that interval." Why necessary? Any person capable of composing (and few are not) a simple song, must be horrified when reading this assertion denying his ability to conceive something melodious—a song—unless he has a conception of a number of mills, cents, or decis of 1/12 log 2. It is no wonder that Helmholtz refused to admit 'additions' to any posthumous edition of his own book. To Helmholtz the unavoidable attempts of piano manufacturers to agree for their practical purposes on an ideal tuning of the eleven piano strings between two octave tones were of *no psychological* interest in his endeavor to establish a scientific theory of music.

Ellis's infatuation with 'cents' has had a remarkably hypnotizing effect

on some investigators occupied with problems relating to acoustics and music, probably thinking that they were inspired by "the great Helmholtz." I shall exemplify: in *Acoustical Terminology,* published by the American Standards Association in 1951, we find on p. 26 a table giving the equivalents of 16 musical cycle-ratios, adding in each case *three decimals* after the 'cent' units. For what purpose? The manufacturer's practical tuner can not comply even with 'cent' units, much less thousands of these units. For whom or for what were the three decimals added? Another table on the same page gives the cycle frequencies for 88 piano keys which in *imagination* are tuned according to the 'cent' system of temperament, one key of the absolute frequency 440,000 [Notice the decimals!] As a matter of fact such a piano true to all those decimals does not exist and never will exist. A psychologist making all these decimals of these 88 frequencies the basis of a psychological theory of music would stultify himself. No such table nor any approach to it is included (for sound reasons) in any edition of Helmholtz's book.

The concept of cents (or micros or mills or decis) has no significance for the psychologist in the study of hearing and musical theory.

Miami, Florida   MAX F. MEYER

THE RADIAL ILLUSION

The illusion described here was first observed by me during the spring of 1953 when I was riding in a car on a dark night during a pouring rain. The only light present was that from the headlights. I was sitting beside the driver. My attention was attracted to the falling rain and, much to my astonishment, I perceived the rain not to be falling vertically but, rather, radiating from a center. The raindrops appeared to be going in all directions (360°) from a point of radiation. I knew all of the rain was falling down, but experientially some of the rain appeared to be going straight up.[1]

The surprising thing about this experience is that I have ridden many times in an automobile at night in a rain and yet I had not before observed this illusion. When once observed, however, it is strong and compelling. After this experience, I took two friends with me on a similar ride and they too verified my observations. I also asked several friends to recall whether they had observed this illusion, and only one said that he had. His recollection, however, was an 'ah haa!' experience. Though he had

[1] For a similar illusion observed under similar conditions, see K. M. Dallenbach, The elastic effect: An optical illusion of expansion, this JOURNAL, 66, 1953, 634-636.

never consciously thought of it before, he gave a description in retrospect of the illusion which validated my own observations. It is my opinion that many have 'seen' this illusion, but few have been consciously aware of it.

I studied the illusion many times after these initial experiences. Various illusory effects appear at different speeds of travel. These are summarized in Fig. 1. The general relations between the center of radiation are shown
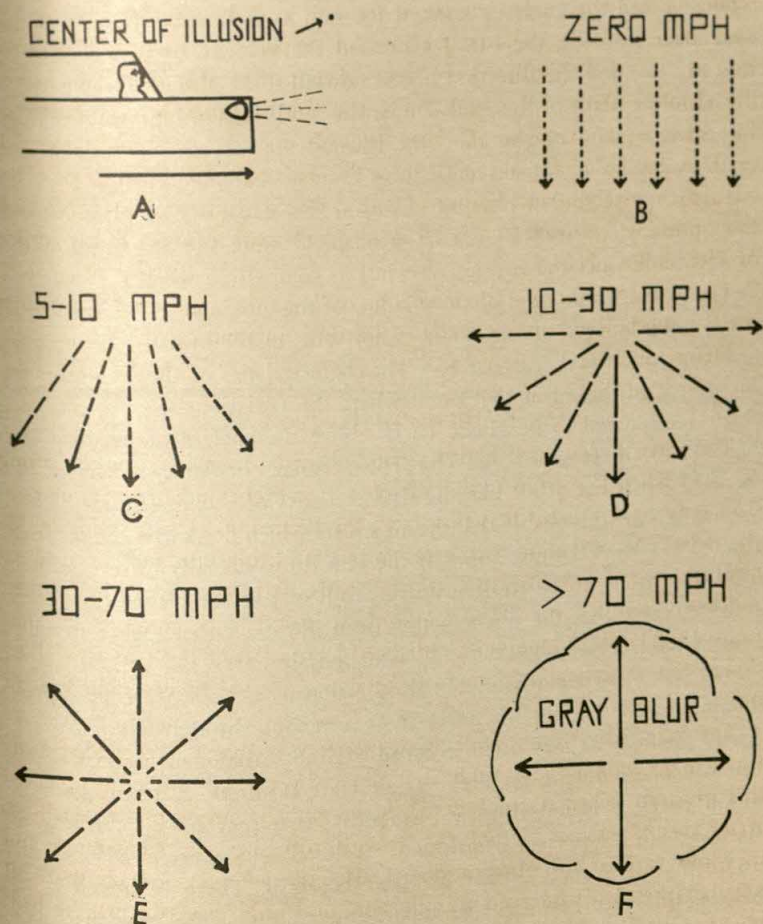
FIG. 1. DIAGRAMS OF THE ILLUSION
A. Position of the center of the illusory radiations; B–F. Apparent path of raindrops at different speeds of movement.

in Fig. 1A. The other diagrams of Fig. 1 illustrate what I saw at different speeds. The perceived motion was perpendicular to my line of sight.

I noted two conditions which modified the illusion. First, the illusion disappears when the headlights of another car approach O and when O's car approaches a well-lighted area, such as a town. Secondly, the center of radiation shifts when the O moves his head from side to side, always in the direction of the motion of O's head.

I had noted the angle or slant of the rain as well as the velocity of the wind after stopping the car. I estimated the velocity of the wind to be close to zero, and the direction of the rainfall to be almost a true perpendicular to the plane of the road. Under the conditions of high wind velocity and extreme slant of the rain, the illusion appears as an irregular and erratic radiation of streaks emanating from a shifting center.

It does not seem that physical factors are the cause of this phenomenon. The optical characteristics of the windshield were not the cause of the phenomenon, since the illusion was just as compelling when I attended to the rain while leaning out of the window of the moving car. The air stream of the outer surface of the car does not seem to be the cause of the phenomenon for several reasons: (a) The characteristics of the air stream are not present in the air at that point where the radiations seem to emanate, which I estimated to be about 12 ft. above the surface of the road and about 12 ft. in front of me. (b) The peculiar circular symmetry of the phenomenon is not at all like the characteristics of the air stream. Rather, it seems highly improbable that the air stream which flows over the car from front to rear would cause some of the rain drops actually to reverse their direction, that is, to go straight up perpendicular to the plane of the road. It seems, then, that the phenomenon is an illusion rather than a straightforward, undistorted observation of actual physical events.

There is no obvious reason why this illusion should not, with appropriate illumination, be present whenever an O is moving through falling objects. Similar phenomena may occur in broad daylight. Analogous illusory effects also may occur during air flight or air craft landings. A recent study in motion parallax found illusory phenomena that may have something in common with the present description.[2] It described the perceived flow pattern of retinal stimuli during air craft landings. There is, however, an obvious difference between that report and my own. The previous paper reports the perceptual effects that occur when O is in motion relative to a

² J. J. Gibson, Paul Olum, and Frank Rosenblatt, Parallax and perspective during aircraft landings, this JOURNAL, 68, 1955, 372-385.

continuous, stable stimulus, *i.e.* the ground. In my report I have described perceptual effects that occur when *O* is in motion relative to discrete, moving stimuli. Further study of these phenomena seems warranted.

Montana State University   FRANK M. DU MAS

THE COMMUNICATION-VALUE OF CONTENT-FREE SPEECH

In a paper delivered in 1953, Soskin presented a view of vocal communication as the simultaneous operation of two communication channels.[1] One of these is the pattern of sound resulting in words, phrases, and other potential 'semantic information'—the content of speech. The second is the voice quality, resulting in potential 'affective information'—in terms of communication the carrier upon which the content is superimposed. The voice quality is also thought to be under less conscious control than the content and to contain information that may be at variance with the content of the message.

That voice quality does carry additional information about the speaker is a matter of common understanding, but it has received little investigation. Early imaginative attempts to isolate quality from speech-content have resulted in artificial situations, such as an interview in which subjects 'talk' using only the alphabet or numbers.[2] A study was designed to investigate the possibilities of another method, adapted from Soskin, of rendering pre-recorded speech content-free.

Method. Tape recordings of voice selections were duplicated on a second tape through a low-pass filter such that frequencies above about 300 ~ were severely attenuated.[3] The voice was reduced to a low mumble and content was lost. Probably a good deal more was lost also, though some indication of pitch, rate, loudness, and their variability remained. Three samples, 30 sec. each, of the speech of Senator McCarthy and Mr. Welch were selected from recordings made during the 1954 Army-McCarthy hearings. They were chosen from partial recordings of three days of the hearings to fit as closely as possible three rough categories of context, named *matter-of-fact, challenging,* and *indignant.*

---

[1] Symposium: Communication in the counselling situation, *Amer. Psychologist,* 8, 1953, 272.

[2] C. W. Thompson and Katherine Bradway, The teaching of psychotherapy through content-free interviews, *J. consult. Psychol.,* 14, 1950, 321-323.

[3] For a description of apparatus designed to isolate non-intelligible vocal components of speech see P. E. Kauffman, An investigation of some psychological stimulus properties of speech behavior, Unpublished Ph.D. dissertation, University of Chicago, 1954.

The six samples were presented twice in a counterbalanced order as though they were 12 different voices. During 10-sec. silences between them, 12 clinical psychologists were asked to judge the filtered samples on five-point scales for 'amount of emotion expressed,' and 'pleasantness.' During a third presentation the judges were given the names used by E to describe the three context categories and asked to choose the most appropriate for each sample. This last step was then repeated with a normal unfiltered recording of the voices. Until this point judges were unable to identify the voices or relate them to the Congressional hearings.

Results. Each rating scale was analyzed separately as a four-variable experimental design, using the variables of voices, contexts, presentations (two for each sample), and judges. In an analysis of variance, all significance tests were made against the third order interaction term.

Table I presents the mean ratings of both presentations of the filtered samples. Looking first at the ratings of amount of emotion, Senator Mc-

TABLE I

RATINGS OF AMOUNT OF EMOTION AND PLEASANTNESS OF FILTERED SAMPLES OF SPEECH
(Mean of two ratings by 12 judges)

	Context		
	matter-of-fact	challenging	indignant
Amount of emotion expressed			
McCarthy	3.17	3.33	3.33
Welch	2.62	2.62	4.21
Pleasantness			
McCarthy	2.58	2.58	2.38
Welch	3.00	3.21	2.25

Carthy's voice was rated close to the same level across the three samples of context, while Mr. Welch was rated below this level for the first two categories and above this for the *indignant* category. There is no overall difference between the voices. There is, however, a large difference between the categories and there is an interaction between voices and categories (both beyond the 1-% level), in keeping with the shift in Welch's level. On the ratings of pleasantness Welch's voice was rated as more pleasant except in the *indignant* category. Differences between voices and between categories are significant (5-% and 1-% level respectively).

There are no significant differences between judges on either rating scale. There was however a shift in ratings from the first to second presentation of the same speech-sample. More emotion was judged present as more judgments were made (first presentation mean = 3.00, second presentation mean = 3.43), and samples were judged slightly more pleasant as more judgments were made (first presentation mean = 2.62, second

presentation mean = 2.71). These differences are significant at the 1-% and 5-% level respectively. Counterbalanced presentation order partially controlled this effect and there are no interactions with it.

Discussion. In spite of the judges' insistence that they had no confidence in their ratings, results are congruent with the hypothesis that some affective information remains in speech filtered content-free.

No claim is made for representative sampling from possible voice qualities that might be emitted by these two men in the three situations. In addition the situations themselves are not directly comparable. In the samples chosen, however, Mr. Welch's voice quality is appropriate to content and Senator McCarthy's is judged without variation. This lack of variation in McCarthy's voice accounts for the judges' inability to agree as to the appropriate context category for his filtered voice, though they did agree when content was present.

University of California   JOHN A. STARKWEATHER

THE EFFECT OF BRIGHTNESS UPON REVERSIBLE PERSPECTIVES AND RETINAL RIVALRY

A previous study from this laboratory in which the Necker cube was used, found that neither the amount of contrast between figure and ground, nor the degree of general illumination, shows an influence upon the rate of fluctuation[1]—a finding at variance with the theory that contrast accounts for the rate of shift.[2] It was suggested in the earlier study that since these sensory conditions are ineffective, and fluctuations are highly irregular, there may be a complicating factor, *e.g.* motor, which masks the effect of processes in the visual center.

In the present study a Gestalt figure-ground design (a simplified Maltese cross) was used,[3] with exactly the same procedure, and under the same conditions as before employed. This figure, a 2-in. black cross with arms (45° in width) radiating from a central point, was pasted on a large white background. The arms, whose outer edges were straight, were then joined by straight lines of India ink, thus forming a white cross similar to the black one, and making the whole design an octagon. This design was duplicated in black and medium gray and placed on a medium gray back-

[1] *Cf.* H. K. Mull, Nancy Ord, and Nan Locke. The effect of two brightness factors upon the rate of fluctuation of reversible perspectives, this JOURNAL, 67, 1954, 341-342.
[2] Wolfgang Köhler, *Dynamics in Psychology*, 1940, 73 ff.
[3] Köhler, *Principles of Gestalt Psychology*, 1935, 191, figure 60 a.

ground. As in the earlier study, the three conditions (A, B, and C), described below, made it possible to study the effect (1) of different degrees of contrast between figure and ground with illumination held constant, and (2) of different amounts of illumination with contrast held constant, upon rate of fluctuation.

Observations were made in a dark room by 36 Os, all young women students between 19-22 yr. of age. Serving individually, each O fixated the center point of the cross and viewed it for 2 min., during which she reported her fluctuations as they occurred. These were timed by E with a stop-watch. To equalize possible temporal effects, the three conditions were presented in counterbalanced order, each of 6 groups of 6 Os being given one of the 6 possible orders of the conditions: *e.g.* A B C, A C B, etc.

In Condition A, *i.e.* black-white design under 200-w. illumination, the average length of the fluctuations was 14.4 sec.; in Condition B, black-gray design, 200-w., 13.7 sec.; in Condition C, black-white figure, 15-w., 13.6 sec. The critical ratio CR between the average of Conditions A and B is 0.28; of Conditions A and C, 0.32.

As an extension of this study, retinal rivalry was investigated under the same three conditions. Squares of saturated red and saturated blue, 2 cm. to the side, were presented stereoscopically on white and on dark gray grounds, red to the right eye, blue to the left. The same 36 Os were instructed to report a change when more than half of the area changed color, and if the square appeared purple, to wait until a predominant color appeared. The following results were obtained: in Condition A the average length of phase was 18 sec.; in Condition B, 15.4 sec.; and in Condition C, 36.5 sec. The CR between the averages of Conditions A and B is 1.1; of Conditions A and C 3.9.

Comparing average fluctuation rates of reversible perspectives with those of retinal rivalry under corresponding conditions, the CR between them for Condition A is 1.1; for Condition B, 1.1; for Condition C, 5.3.

It appears then: (1) that neither in reversible perspectives nor in retinal rivalry has contrast a noticeable effect on rate of shift; (2) that in reversible perspectives the degree of general illumination is not noticeably effective; (3) either the rates of shift in reversible perspectives and in retinal rivalry are not reliably different in bright light—regardless of contrast or color; or the rate of shift between colors in retinal rivalry is markedly slower in dim than in bright light; and (4) that all rates are variable for all Os under all conditions.

The results of the earlier study are confirmed by these data, and the explanation suggested above is again offered. In addition, the shifts of

retinal rivalry for colors seen in bright light may be supposed to involve much the same mechanism as that accounting for the reversals of perspective, since the fluctuation rates are so much the same for both. There is one exception, however: in faint light, the rate of rivalry shift of colors is much slower. In this case one might suppose that the retinal color processes are barely able to be aroused, and as a result of such weak stimulation, motor responses are accomplished more slowly.

Sweet Briar College.

HELEN K. MULL
GRETCHEN ARMSTRONG
BARBARA TELFER

WORD-ASSOCIATION AND WORD-FREQUENCY

If we ask what determines the response-word in free association, one obvious answer is S's vocabulary. The words that occur most often must be the words with which S is most familiar. The *Teacher's Word Book* by Thorndike and Lorge gives the frequencies of use of 30,000 words in standard reading matter.[1] The Kent-Rosanoff word-association tables give the frequencies of the response-words of 1000 Ss to 100 common stimulus-words.[2] Do the most frequent response-words in the Kent-Rosanoff list have high Thorndike-Lorge frequencies?

Ten of the K-R tables (Numbers 1, 11, 21-91) were examined to find the ten most frequent response-words in each table. Ten of the least frequent response-words also were selected at random from each table, the frequency being *one* in all cases. Thus we have the hundred most common response-words to be compared with a hundred of the least common response-words. The T-L tables tell us which of these are high-frequency words, that is, which occur 50 or more times per million running words. Making the appropriate tabulations, we find that 84% of the most common K-R response-words have high T-L frequencies, as compared with only 48% of the least common K-R response-words.

Other comparisons are equally impressive. Let us assume that the K-R Ss could have produced any of the 30,000 words in the T-L word-book. Of these 30,000 words, only 2,021 or 7% are high-frequency words. In comparison with all 30,000 words, therefore, even the least common K-R words are distinctly common. If we take the English language of some

[1] E. L. Thorndike and I. Lorge. *The Teacher's Word Book of 30,000 Words*, 1944.
[2] A. J. Rosanoff, *Manual of Psychiatry*, 1927, 552-600.

500,000 words as a base, the 2,021 high-frequency words constitute only 0.4%. These comparisons permit the statement that words spoken in response to *K-R* stimuli are far more familiar than words not spoken. There is a possibility that these differences include a difference between spoken-word frequencies and written-word frequencies, but we have no precise data on this point.

As a check on these results, similar comparisons were made with the data arranged to permit rough calculation of medians. The 650 response-words of the first five *K-R* tables were sorted into three groups according to frequency: a group of 360 that appeared only once each, a group of 138 that appeared two or three times each, and a group of 152 that appeared four or more times.

The median *T-L* frequency of the words with *K-R* frequencies of *one* is 42 per million. The median *T-L* frequency of the words with *K-R* frequencies of *two* and *three* is 65. The corresponding median for the words with *K-R* frequencies of *four* or more cannot be calculated but it must be over 100 since more than half these words have *T-L* frequencies over 100. Even the words with the lowest *K-R* frequencies are relatively common as compared with the entire sample of 30,000 words in the *T-L* book, the median *T-L* frequency for which is roughly two per million.

In an effort to determine the magnitude of the relation between *K-R* frequency and *T-L* frequency, several contingency tables were arranged in various groups, using the 650 words of the first five *K-R* tables. In no case could a significant association be demonstrated because of the great variability within each array. Nearly all words of high *K-R* frequency have high *T-L* frequencies, but many words of low *K-R* frequency also have high *T-L* frequencies. Other variables that influence the response-word, in addition to the familiarity of the possible response-words, are, of course, the stimulus-word and the set.

In general, we can say that the words given as free associations to common stimulus-words are among the common words of our language. They are among the more common words of the Thorndike-Lorge *Word Book*. Presumably they are produced more often because the availability of a word for free association depends in part on its general familiarity or response-strength regardless of any associations with specific stimulus-words. When the most frequent *K-R* response-words are compared with less frequent response-words, same relationship to general familiarity appears, but it is a weak relationship that appears only when the data are

favorably grouped in broad categories. No one would claim that familiarity is the only determinant of the response.

Michigan State University   DONALD M. JOHNSON

AVOIDANT VS. UNAVOIDANT CONDITIONING AND PARTIAL REINFORCEMENT IN RUSSIAN LABORATORIES

Both avoidant and unavoidant conditioning were studied in Bekhterev's 'association-reflex' laboratory from its beginning in 1907. Protopopov in his pioneer experiment of shock-conditioning used unavoidable shock with five dogs,[1] whereas Molotkov in conditioning four human adults used avoidable shock.[2] This procedural difference between conditioning dogs and humans was continued for a number of years although in a strict sense a form of unavoidable shock, terminable but unpreventable, was also used by the Russians with human Ss. Their early experimenters favored simultaneous presentations of conditioned and unconditioned stimuli (CS–US).

In 1926, Starytzin introduced avoidable shock with dogs. He compared the conditionability of dogs under avoidable and unavoidable shock.[3] His results antedate those of Schlosberg's[4] with rats by 8 yr. hence are worthy of reporting. Starytzin found avoidable shock superior to unavoidable. The conditioned flexor responses (CRs) of his dogs (3 in number) became stabilized under avoidable shock in 28 to 50 reinforcements; under unavoidable shock one dog required 150 reinforcements to stabilize and the other two did not develop a stable CR after 320 reinforcements in 8 experimental sessions.

Starytzin's procedure was modified and perfected by Petropavlovsky,[5] and since then Russian experimenters have been using both avoidable and unavoidable shock with dogs and other animals. Despite the fact that the

---

[1] W. Protopopov, [*Motor Association-Reflexes to Sound*], Thesis, St. Petersburg, 1909.

[2] A. Molotkov, [*The Formation of Motor Association-Reflexes to Visual Stimuli*], Thesis, St. Petersburg, 1910.

[3] S. E. Starytzin, [The method of forming motor association-reflexes in dogs through stimulation of the pads of their paws], in *Sbornik Posviashehonny Bekhterevu*, 1926, 133-145. As far as the writer is aware, this historically important study has never been abstracted—or mentioned—in any non-Russian language. It also missed being included in the writer's lengthy bibliography of classical conditioning (*Psychol. Bull.*, 34, 1937, 191-256).

[4] Harold Schlosberg, Conditioned responses in the white rat, *J. genet. Psychol.*, 45, 1934, 303-335; 49, 1936, 107-138.

[5] V. P. Petropavlovsky, [The methodology of conditioning motor reflexes], *Fiziol. Zh. SSSR*, 17, 1934, 217-225.

Russians have as a rule found avoidable shock more efficacious,[6] they have not conceptualized it as a distinct form of conditioning. Conceptualization of avoidable shock began in this country with Schlosberg, who found avoidable shock, if anything, inferior to unavoidable shock.[7] All the Russians seem to say about the superiority of avoidant conditioning is that it is "more natural," more "in line with animal evolution," that it results from partial reinforcement, or, more exactly, from 'non-overreinforcement.'[8]

The last statement merits special attention. Some years ago, reviewing critically Russian experimental literature on the conditioned response, the writer expressed on several occasions a conviction that overreinforcement operates as an unmistakable negative factor in conditioning, especially in shock-conditioning.[9] An analysis of more recent Russian studies supports fully that conviction. Indeed, notwithstanding some contrary evidence by Gibson and by Solomon and Wynne,[10] one might put up a good argument for the contention that whatever qualitative distinctness and superiority avoidant conditioning manifests over unavoidant conditioning may well be but a case of the qualitative distinctness and superiority of partial reinforcement in conditioning, and that, moreover, the superiority of partial conditioning itself may in a large part be only a matter of 'non-over-reinforcement.'

With respect to partial reinforcement, the writer is very much impressed by Russian evidence on the very rapid acceleration of the CR rates of reconditioning and recurrent reconditioning. Several post-extinction reinforcements usually undo the negative effects of hundreds of preceding nonreinforcements and often raise CRs to considerably higher levels of magnitude.[11] It is not, then, logical to assume that the CR superiority of

[6] L. S. Gambaryan, [The problem of conditioned defense reflexes], *Trud. Inst. Fiziol. Pavlova*, 1, 1952, 73-84.

[7] Schlosberg, The relationship between success and the laws of conditioning, *Psychol. Rev.*, 44, 1937, 379-394.

[8] Gambaryan, *op. cit.*; D. Biryukov, The comparative physiology and pathology of conditioned reflexes, [*Trans. 15th Conf. Higher Nerv. Activ.,*] 1952, 166-183.

[9] G. H. S. Razran, Theory of conditioning and related phenomena, *Psychol. Rev.*, 37, 1930, 25-43; Conditioned responses in children, *Arch. Psychol.*, 28, 1933 (No. 148), 1-Conditioned withdrawal responses with shock as the conditioning stimulus in adult human subjects, *Psychol. Bull.*, 31, 1934, 111-143; The law of effect or the law of qualitative conditioning, *Psychol. Rev.*, 46, 1939, 445-563.

[10] E. J. Gibson, The role of shock in reinforcement, *J. comp. physiol. Psychol.*, 45, 1952, 18-30; R. L. Solomon and L. C. Wynne, Traumatic avoidance learning: the principles of anxiety conservation and partial irreversibility, *Psychol. Rev.*, 61, 1954, 353-385.

[11] I. P. Pavlov, [*Pavlov's Wednesday Seminars*], 3 vols., 1949; G. Swejkowska, Chronic extinction and restoration of conditioned reflexes: V. Repeated extinction

partial reinforcement is, besides being a function of non-overreinforcement, also a function of the superior *CR* rates of 'recurrent' as compared with 'precurrent' conditioning? The writer offered elsewhere some experimental evidence for the tenability of this hypothesis and, while the evidence may not be wholly conclusive, it certainly suggests that delimitations of optimal maxima of reinforcements and of parameters of rates of recurrent re-conditionings would go a long way toward settling a number of current controversies.[12]

Queens College
GREGORY RAZRAN

THE ONE-HUNDRED AND TWENTY-SECOND MEETING OF THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

At the 122nd meeting of the AAAS, held in Atlanta, Georgia, December 26-31, Section I (Psychology) presented a program of five sessions of invited papers and two of submitted papers. All sessions were co-sponsored by the Southern Society for Philosophy and Psychology.

The American Physiological Society co-sponsored two groups of invited papers. One on "Sensory processes" was arranged by John F. Hahn and a second on "Brain function" was arranged by Harlow W. Ades. Other sessions of invited papers on the topics of "Military psychology, learning, and primate behavior" were arranged by Wilse B. Webb, Stanford C. Ericksen, and Arthur J. Riopelle. Section I co-sponsored, with Section L (Education), a symposium of "Creativity in science."

Clarence H. Graham (Columbia University) was elected Vice-President and Chairman of Section I for 1956, and Conrad G. Mueller (Columbia) was elected to succeed W. D. Neff (Chicago) as Secretary and Frank J. Geldard (Virginia) to succeed Roger G. Barker (Kansas) as Committee-man-at-Large.

University of Chicago
W. D. NEFF

THIRD INTER-AMERICAN CONGRESS OF PSYCHOLOGY

The third Inter-American Congress of Psychology, sponsored by the Inter-American Society of Psychology, was held at The University of Texas,

---

[12] and restoration of conditioned reflexes, *Acta Biol. Exp.*, 16, 1952, 115-122; *Trudy fiziologicheskikh laboratorii Akademika I. P. Pavlova*, 12 vols., 1926-1945.

[12] Razran, Partial reinforcement of salivary conditioning in adult human *Ss.*, *J. exp. Psychol.*, (in press).

Austin, Texas on December 16-20, 1955. The central theme of the Congress, "The psychology of social tensions," was discussed in the following panels: "The human relations laboratory for investigations of social tensions," "Tension-mapping the psycho-social world of the school system," "Approaches to inter-group and international understanding," "Intercultural tensions in Europe, the Near East, and North America," "Intercultural tensions in Mexico and the United States," and "Tensions in childhood." The papers presented in the general sessions dealt with the problems of tension in communities and in psychotherapy.

In addition to the United States, delegates were in attendance from Argentina, Canada, the Dominican Republic, Jamaica, Mexico, Panama, and Venezuela. The Congress was supported by The Anti-Defamation League of B'nai B'rith, The Hogg Foundation for Mental Hygiene, The University of Texas, and the State Department of the United States. At the opening banquet, Professor F. L. Cox welcomed the delegates to the University on behalf of the President, and Dr. L. Manuel Espinosa, Chief of Professional Activities of the State Department's International Exchange Program, gave the keynote address.

Officers of the Inter-American Society of Psychology, elected for 1956, are: President, Otto Klineberg, Columbia University, U.S.A.; Vice-President, Guillermo Davila, M.D., National University, Mexico; Secretary General; Werner Wolff, Bard College, U.S.A.; Treasurer, Gustave M. Gilbert, Michigan State University, U.S.A.

The papers of the Congress, given in Spanish or English, were translated, as they were read, into English or Spanish. The translations were led by a telephonic system to earphones worn by members of the audience. Papers in either of these official languages could, therefore, be understood by all.

The 1956 meeting of the Society will be held according to present plans, in December in the Caribbean area.

Bard College

WERNER WOLFF

Louis Leon Thurstone: 1887-1955

Thurstone first became a reality for me in the summer of 1930. As a graduate student impressed by the good sense of the German custom of sampling several universities, I spent two quarters at the University of Chicago and while there registered for one of Thurstone's courses. It was an excellent time to get acquainted, for the ideas that occupied most of Thurstone's professional life were all bubbling at once. In one summer we covered psychophysical measurement, in which most of his important papers had recently appeared or were being thought out; attitude measurement, for which the basic methodology had recently been worked out and some of the individual scales developed; prediction of choice, which he was thinking about but on which his major publications did not appear until considerably later; the rational equation of the learning function, to which his interest had returned after a lapse of a dozen years; and factor analysis, on which his first paper appeared a year later. Such a wealth of original ideas!

Thurstone's contributions to psychology have consisted primarily of better tools—powerful analytic methods and techniques of measurement of widespread usefulness. In recent years he has been so closely identified with multiple factor analysis that it is easy to forget his earlier contributions. Yet they continue to be used; among all psychologists he is one of the leaders in terms of frequency of citation.

Thurstone was forever interested in applying psychological methods to the solution of practical problems, not isolated and particular problems, but classes of problems of practical significance. As a graduate assistant in Walter Bingham's original Division of Applied Psychology at Carnegie Institute of Technology, in World Wars I and II, at the Institute for Government Research in Washington, D.C., at the Institute for Juvenile Research in Chicago, and throughout three decades at the University of Chicago, there was a constant turning to problems of practical importance. It was impatience with the triviality of earlier psychophysical experiments on lifted weights and difference limens that led him to develop the psychophysics of attitudes, judgment, esthetics, choice and other 'real' problems. His background in machine design made him one of the first psychologists to be concerned with what has since come to be called the field of human engineering. His work on mental measurement ranged over tests for college admission, radio-code-learning ability, mechanical aptitude, personality and attitude scales, and an activity in which he took particular pride—establishment and direction of the Board of Examinations at the University of Chicago.

This underlying interest in practical psychological problems contrasted strangely with his reputation among persons who did not know him well. To them he frequently seemed completely absorbed in the more esoteric realms of factorial methods; but factor analysis to Thurston was a tool, the mathematics an aid to psychology, never an end in itself. If the results of a factorial study did not make good psychological sense, the analysis was either incomplete or of no value. This insistence was basic to many of his arguments with others interested in factorial methods. Most of those arguments were concerned with the methods of factor analysis or with the philosophical reality of primary mental abilities. These were important topics in a number of papers and in *Vectors of Mind* (1935) and *Multiple-Factor Analysis* (1947). Differences of opinion on these points have somewhat masked the fact that factor analysis as a practical tool produced major changes in aptitude and ability testing for educational, industrial, and military purposes.

In personality, too, there were contrasts. To those who did not know him well, Thurstone often seemed aloof and unapproachable. He sometimes unintentionally fed this reputation by expecting too much of his audience. I recall hearing him present a paper which began with an essentially simple equation that was nevertheless long enough to appear forbidding to the uninitiated. Before he had an opportunity to explain, he was interrupted by laughter; he had lost his audience. He could also be impatient with muddled thinking and pretentiousness, and sometimes criticized others with a dogmatic assertiveness that contrasted sharply with the modesty he showed toward anyone who was seriously trying to learn or understand. His autobiography in Volume 4 of the *History of Psychology in Autobiography* opens with this disclaimer: "The biography of an individual scientist cannot be expected to be of general interest except when there has been a spectacular achievement or a colorful personality or both. The present case has no claim to either." Later there are such statements as "I have not been productive in that field." He once told me that he was forced to work on psychological problems that others had not tackled because he had such a poor memory he could not remember what had been done before. When I first knew him he was taking correspondence courses in mathematics, and for several years he engaged a tutor to help improve what he considered his inadequate knowledge of that field.

Aloof he may have been to strangers, but to his colleagues and to those who were interested in getting at the roots of a problem, he was

approachable, a willing helper, and a pleasant companion. He was at his best when working with a group on a difficult problem. Chicago colleagues with kindred interests had weekly opportunities to work with him in such a setting. In 1936 I joined the staff of the University of Chicago and from then until World War II was a regular member of the weekly seminar group that met at his home. He and Mrs. Thurstone, in their large comfortable study with its much used blackboard, and with the coffee and Swedish pastry that ended the formal two-hour seminar at 9:30 but never ended the talk, provided the setting for sessions that were educational experiences for all of us, Thurstone included. They were a testing ground for the ideas that came to any of us, for we could speculate and suggest and criticize, certain that someone would see new implications and that someone else would spot the loopholes in whatever we were trying to develop. The group included his graduate students and others from his own staff, from the Board of Examinations, and from other departments or institutions. There were, I remember, Nicholas Rashevsky and Herbert Landahl from mathematical biophysics, Gale Young and Alston Householder, then from mathematical biophysics but now better known in the nuclear reactor field, Gerhard von Bonin from neuroanatomy, Samuel Stouffer from sociology and Karl Holzinger from education, whose divergent ideas about factor analysis always meant that problems in that area were looked at from more than one point of view.

Thurstone's life was a wide-ranging one. Born in Chicago in 1887, his boyhood saw a series of moves, to Mississippi, to Sweden, to New York. In his undergraduate days he switched from civil to electrical to mechanical engineering. His first development was a motion picture projector that eliminated the bothersome flicker of projectors then in use. A demonstration to Thomas Edison brought praise for a better projector, the decision not to produce it because of the necessity of completely retooling the Edison factory, and the offer of an assistantship to Edison. From Edison's laboratory Thurstone went to the University of Minnesota to teach engineering, and left there for graduate work in psychology, prompted to do so by an interest in quantitative analysis of learning. His Ph.D. was received from the University of Chicago in 1917.

From that time on, his basic interest in quantifying psychological problems was evident in all his work. He remained at Carnegie Institute of Technology from 1915 to 1923, with time out to serve as a member of the trade-test division of the psychological staff of the Adjutant General's Office. In 1923 he moved to Washington, D.C., to the newly established

Institute for Government Research where he developed material for civil service examinations. In 1923 came the move to Chicago, in 1927 promotion to a professorship, and in 1938 promotion to the Charles F. Grey Distinguished Service Professorship at Chicago. Other honors included the presidencies of the Psychometric Society in 1936, the Midwestern Psychological Association in 1930, and the American Psychological Association in 1932; election to the National Academy of Sciences, the American Philosophical Society, and the American Academy of Arts and Sciences. When he retired from Chicago in 1953, several universities offered him appointments. He chose the University of North Carolina, to which the Psychometric Laboratory was moved. Mrs. Thurstone, who had been a close collaborator in most of his work at Chicago, resigned her position as Director of the Division of Child Study of the Chicago school system to accept a professorship of education at North Carolina. Thurstone was born on May 29, 1887. He died on September 29, 1955.

Thurstone's pre-psychological interests remained alive to show up in a variety of ways. Photography, at which he was an expert, was a principal hobby. Some of the things he learned in machine design colored later work and probably influenced his interest in analyzing mechanical ability. In talking about creativity—a topic of persisting interest—there were examples and ideas that stemmed from his experience with Edison. His ideas ranged over a wider territory than his experience and work. Lunch and dinner table conversations sometimes explored ideas far removed from psychological measurement, a problem in biochemistry one day, how to teach good handwriting another—his own hand, learned in Swedish schools, was excellent and he was impatient with the ill-formed letters of most of us. He toyed with the idea of preparing instructional material for teaching handwriting; I recall a discussion of whether it would be better received if called the *Thurstone* method or the *Swedish* method. Such topics were always extras to the persisting interest in developing tools to make psychology into *a quantitative, rational science*. That phrase, descriptive of the purpose of the Psychometric Society and its journal *Psychometrika,* of both of which he was the principal founder, describes what he tried to do for psychology. If he had added a third adjective, I think it would have been *practical*.

Several years ago on a visit to Berkeley I was at dinner with a group of the senior psychologists from the University of California. Someone started the game of naming the half dozen psychologists of the world whose work had shown the most ground-breaking originality. Thurstone, we agreed, belonged in that select company.

A.A.A.S., Washington, D.C.

DAEL WOLFLE

William John Crozier: 1892-1955

On November 2, 1955, William J. Crozier died of a heart attack at his home in Belmont, Massachusetts, at the age of sixty-three. Although primarily a physiologist he influenced experimental psychology by his theoretical and experimental contributions to the study of animal behavior and of sensory processes and also as a teacher of graduate students.

Crozier was born in New York City on May 24, 1892. He was graduated from the College of the City of New York in 1912 where as an undergraduate he became interested in physical chemistry and biochemistry. Later, as a zoölogist, these expanding interests led him to the study of quantitative aspects of animal behavior in relation to underlying physiological mechanisms and aspects of chemical kinetics. He received his Ph.D. in zoölogy from Harvard in 1915 and spent his next three years as research naturalist at the Bermuda Biological Station where he made extensive observations and published a number of papers dealing with the behavior of marine organisms. It was during this period that his classic investigations on chemo-reception were carried out. The year 1918-19 he spent as assistant professor of physiology at the University of Illinois Medical School, and during the following year he was assistant professor of zoölogy at the University of Chicago. In 1920 he was called to Rutgers, University as professor and head of the department of zoölogy, where he remained until he joined the Harvard faculty in 1925 as associate professor and head of the newly formed Department of General Physiology. In 1927 at the early age of thirty-five years he was made full professor at Harvard.

His laboratory of general physiology in the late twenties was an exciting place for graduate students. Crozier was a stimulating and provocative teacher who was not only remarkably learned but who gave his students a sense of participation in advancing ideas and of sharing in the adventure of cross-discipline research. So it was that he early found himself surrounded by an enthusiastic group of students and research fellows nearly all of whom, men now in their fifties, hold professorships or other responsible position in physiology, biochemistry, biophysics and psychology.

At Rutgers and at Harvard, Crozier and his students divided their research between two major fields during the twenties and early thirties.

The first of these fields was concerned with tropisms, the orientation of intact organisms in fields of force. Jaques Loeb, whose work was of much interest to Crozier, had earlier investigated a number of tropistic phenomena in animals: phototropism, geotropism and chemotropism. Loeb's many contributions raised challenging questions which Crozier attacked by refinements of experimental technique and the application of ingenious

mathematical analyses of the pathways taken by animals in response to light and to gravity. Thus, on the assumption that organisms orient in a light field and take a path such that illumination on the two photo-receptors becomes equalized, Crozier worked out equations that could be experimentally tested to confirm this hypothesis. He further tested the assumption that organisms orient in a gravitational field until the balance of activated receptors on the two sides of the body is equalized. Quantitative formulations of the gravitational vectors acting upon animals moving on inclined planes not only confirmed this general hypothesis but also made possible statistical treatments of the nature of excitation of groups of receptor units according to their thresholds of activation.

Crozier and his students further showed that the levels of excitability of such receptor groups are genetically determined. They studied highly inbred strains of rats, determining significant parameters in the equations that described the orientation of their creeping on inclined planes. Back-crossing such strains produced values of behavior constants that predicted the geotropic conduct of the rats in terms of genetic factors. The general equations for geotropism were found applicable to a variety of anatomical forms including bipeds (chicks), quadrupeds (rats and mice), gastropods (snails), and insects with multiple pedal appendages. By careful experimental control Crozier found it possible to balance effects of phototropism and geotropism so as to produce predictable behavior on the part of orienting organisms.

The second major problem occupying Crozier's interests at this same time involved analyses of the effects of temperature on biological oxidations and on rhythms of action in poikilothermous organisms. He and his students demonstrated that remarkably good fits of the Arrhenius equation were found to describe many biological oxidations directly measured in terms of CO_2-production and O_2-consumption. This equation also described many examples of frequencies of movements of heart beats, of breathing movements, of cilia beats, of pedal movements, as well as a variety of other rhythmic activities determined for the most part by nervous mechanisms. The mere fit of the Arrhenius equation to these data was not the significant thing but rather the fact that a constant in the equation, known to physical chemists as the energy of activation, was found to have a limited range of values and to recur again and again in hundreds of investigations of temperature effects on cell respiration and upon biological rhythms. Crozier interpreted these modal values of activation energies, which he called *temperature characteristics*, as corresponding to a limited

number of enzyme systems. These enzyme systems he argued, operate in such a way that in some cell systems one enzyme system and in other organisms (or in the same organism under different conditions) another would act as a rate-limiting bottleneck or pacemaker for the complex sequence of steps determining oxygen uptake, CO_2-production, or energy release that sets basic frequencies of rhythms of discharge in the nervous system. This hypothesis has subsequently been confirmed by other investigators and extended to mammals including man, and the temperature characteristic has proved to be an interesting though limited tool in elucidating controlling chemical kinetics underlying rhythmic events in intact organisms.

Crozier's analyses of behavioral phenomena by carefully controlled experimental design made possible the application of a rational form of mathematical analysis of animal behavior resulting in deeper biological insights than could be obtained with more conventional methods. In his later years he felt his major scientific contribution was in the approach which he called *parametric analysis*. Parametric analysis is not to be confused with mere fitting of rational equations or the description of data by empirical equations, but rather is described as the identification and study of meaningful parameters in equations describing aspects of behavior in such a way as to furnish tests for hypotheses concerning underlying mechanisms. This approach was particularly evident in Crozier's research on visual processes, carried on in recent years.

In the mid-1930's the Harvard Department of General Physiology was abolished when the three departments of the Division of Biology were reorganized into a single Department of Biology, with physiology joining with zoölogy and botany in a single large department. At this time Crozier was made Research Professor of General Physiology without formal teaching duties, and it was after this change that the study of visual mechanism became his third major object of attention.

In collaboration with Ernst Wolf and others he carried out a great number of experiments of basic importance to the understanding of the phenomena of critical flicker-frequency. The effect of some 20 variables upon the flicker threshold were studied and many original discoveries were made. A recurrent theme in this work was emphasis upon the argument that understanding of visual phenomena is not given by the photochemical properties of the retina alone but must take cognizance of the entire visual nervous system. More recently Crozier's interest turned to research on threshold responses of the human visual system, with particular reference

to the effect of wavelength as the independent variable. Some of this work is reported in his last published paper.[1] This paper is particularly interesting to those who would know Crozier as it gives again many of his theses concerning the nature of the visual processes and is marked by the typical flavor of his style of exposition.

During the war Crozier served with the Air Force in the Pacific Theatre as Operations Analyst with the equivalent rank of colonel. Afterwards he returned to Harvard and continued an active program of research until his death. He was the author of approximately 300 papers. He published for the most part in the *Journal of General Physiology* of which he had served as an editor since 1924. At the time of his death five partially completed manuscripts were found. Two of his recent Ph.D. students hope to be able to complete the preparation of at least one of these for publication. Crozier had refrained from publication during the last four years. He had been accumulating much data and expected to write a monograph on his visual work.[2] His loss to science is thus especially great since so much of his important experimental work of recent years must remain unfinished.

Worcester Foundation for Experimental Biology,   HUDSON HOAGLAND
Shrewsbury, Mass.
Massachusetts Institute of Technology   R. T. MITCHELL

---

[1] W. J. Crozier, On the visibility of radiation at the human fovea, *J. gen. Physiol.,* 34, 1950, 87-136.

[2] Two reviews of Crozier's earlier work, together with extensive citations to his papers, are contained in *The Handbook of General Experimental Psychology,* ed. by Carl Murchison, 1934. Chapter 1, by Crozier and Hoagland, reviews the work on tropisms and some of the earlier work on temperature characteristics. Chapter 19 by Crozier reviews his work on chemoreception. This book is now out of print but available in psychology libraries. Crozier's studies on vision have appeared for the most part in the *Journal of General Physiology* between 1935 and 1950. An unusually penetrating analysis of nerve excitability is to be found in one of his contributions entitled: Strength-duration curves and the theory of electrical excitation, *Proc. Nat. Acad. Sci.,* 23, 1937, 71-78.

Géza Révész: 1878-1955

Three of the pioneers of European psychology, who became linked in friendship at the beginning of the century in G. E. Müller's Laboratory, have died within the last two years: David Katz, Ph.D. Göttingen (1906), aged sixty-nine, on February 7, 1953; Gustav Kafka, aged seventy, a few days later on February 12, 1953; and now Géza Révész, Ph.D. Göttingen (1905), aged seventy-seven, on August 19, 1955.[1] Like Katz, Révész died suddenly of heart failure. Katz had contributed to the jubilee volume dedicated to Révész in 1950, and Révész to the volume dedicated to Katz in 1951. Katz was obliged to leave his native Germany, with the coming of power of the Hitlerian dictatorship, and, having become a Swedish citizen and professor at the University of Stockholm, he founded a very active Psychological Institute which he had given up, just before his death, because of his retirement. Révész had to leave his native Hungary, at the time of Horthy's *coup de force* in 1919, and, having become a Dutch citizen and professor at the University of Amsterdam, he too founded a Psychological Institute—probably the largest in Europe, with its forty rooms and an auditorium—which, just before his death, he had left on becoming *emeritus*.

An astonishing parallelism of two destinies, reinforced by close friendship and repeated collaborations: research on vision and memory in hens in 1907, and on the vision of nocturnal birds in 1913; investigation of musical capacities in the deaf, based on the vibratory sense, in 1926; culmination of the work each had devoted to touch (Katz's *Der Aufbau der Tastwelt* in 1925, Révész's *Die Formenwelt des Tastsinnes* in 1938); study of the problems of comparative psychology and of animal sociology, with a joint article in the *Zeitschrift für angewandte Psychologie* in 1921 (*Experimentelle Studien zur vergleichende Psychologie*) that led to their separate reports at a symposium of the International Congress for Psychology in Paris in 1937. Here Révész examined "die soziobiologische Funktion der menschlichen und tierischen Hand," and Katz "die Bedeutung der Tierpsychologie für die menschliche Psychologie."

Géza Révész was born December 9, 1878, in Hungary at Siofek, where his father had famous vineyards. He studied law, with an interruption for his military service, at the University of Budapest, before going to Göttingen, which he left for one year to study under Stumpf, Engelmann, and

[1] This biography was translated from the French by Mollie D. Boring.

Du Bois-Reymond at Berlin. Appointed in 1908 *privat Dozent* at the University of Budapest, he was, in 1914, mobilized as an officer in the artillery and, in 1917, as associate professor, he was put in charge of problems of military selection. Becoming full professor in 1918, he went to Holland in 1919, where he was received by Heymans at Groningen and by Zwaardemaker at Utrecht. He soon established a private laboratory at Amsterdam and from 1923 on gave lectures in industrial psychology at the University. When, in 1932, the University created a chair in psychology, it called him to fill it as associate professor. Becoming full professor, he founded, in 1939, the Psychological Institute, of which he was director.

In 1935 he and Katz founded jointly the *Acta Psychologica,* an international journal, which became Dutch-Scandinavian in 1939 and, extended to include Belgium and Switzerland, became 'European' in 1955. Révész has been editor-in-chief of the *Nederlandsche Tijdschrift voor Psychologie,* and he had founded a Dutch Institute of Applied Psychologists.

The work of Révész has been extremely varied, as is shown clearly in the bibliography published in the jubilee volume of *Acta Psychologica,* which lists his work under eleven separate headings: general psychology; psychological optics; psychology of hearing and of music; problems of space, language and thought; talent; child psychology; animal psychology; social psychology; medical psychology; pedagogical psychology. The diversity of his interests is evident if we place his study of Leibnitz's psychology (1917) along side of his report to the International Conference on Psychotechnology in Paris (1928), *Die Rationalization der Packarbeit.*

Without listing the twenty books for which we are indebted to him and the more than a hundred articles in Hungarian, Dutch, German, English and French in numerous periodicals, we must single out the following: *Grundlegung der Tonpsychologie* (1912); *Psychology of a Musical Prodigy* (1925); *Einführung in die Musikpsychologie* (1946); *Die menschliche Hand* (1944); *Ursprung und Vorgeschichte der Sprache* (1946); *Talent und Genie* (1951); *The Psychology and Art of the Blind* (1950); *The Personal and Social Life of the Blind* (1955).

Révész was an original and independent spirit, little given to direct social contacts, a man who above all influenced his pupils through his utter devotion to science. He attended but few congresses. It is because he participated in the one at Edinburgh in 1948 that he could be included in the Executive Committee which founded the International Union of Scientific Psychology. In 1948 he was made a member of the Hungarian Academy of Sciences, and in 1949 he became Doctor *honoris causa* at the Uni-

versity of Würzburg. He has always taken great pains not to separate application from theoretical psychology. Fond of art and particularly of music—happy to be able to play a Mozart sonata from memory on a bad piano that Kafka, who tells the tale, had in his room in Göttingen—he sought particularly in his work to do justice to innate talent and culture in musical creation, requiring the little musical prodigy whom he was studying to perform, between the ages of six and thirteen years, successive improvisations on the same theme. In this way he combined experimentation with the biographical analysis of genius. He believed, moreover, he had firmly established a duality in what is called absolute pitch.

Révész was indeed one of the notable psychologists of the past half-century.

University of Paris   HENRY PIÉRON

ERRATUM

The date of Walter Dill Scott's death was incorrectly given in the necrology published in the last number of this JOURNAL (Vol. 68, 1956, 682-683). He died, according to a letter received from his son, John M. Scott, in answer to our inquiry, at about 10 P.M., September 23, 1955 (not September 24).

Though care is taken in regard to dates, errors like this one do at times occur. They result, as Boring has pointed out,[1] from the fact that it is much more difficult to obtain accurate information about date of death, than about date of birth which is easily obtained because biographical sources listing the living are checked by the living. The necrologizer has, on the other hand, to trust casual sources: a newspaper clipping, a hastily written letter, or a note in another journal. In the present instance, the date reported by another journal was trusted; the newspaper was not; but the newspaper was correct.   K. M. D.

AN ACKNOWLEDGMENT

The JOURNAL is indebted to Mrs. Thelma Gwinn Thurstone for the photograph and signature of her husband that are reproduced in the frontispiece of this number. The photograph was taken by Harris and Ewing, Chicago, Illinois, during the summer of 1952 when Professor Thurstone was 65 years old. The signature, a facsimile of which appears under the portrait, was written during the summer of 1955.   K. M. D.

---

[1] E. G. Boring, Psychological necrology (1903–1927), *Psychol Bull.*, 25, 1928, 302-305, 621-625; Suzanne Bennett and E. G. Boring, Psychological necrology (1928–1952), *ibid.*, 51, 1954, 75-81.

STYLE SHEET
OF
THE AMERICAN JOURNAL OF PSYCHOLOGY

The Style Sheet, published here for the information and guidance of contributors, is a summary of present usage in this JOURNAL. It is the product over the years of many minds—the present editors' inheritance from their predecessors. Changes in style have been made in the past and they will unquestionably be made, in keeping with the advance of the times, in the future, but for the present this Style Sheet will govern the preparation of copy and the handling of proof of the JOURNAL. Authors submitting manuscripts are requested to follow it. It makes no pretense of completeness; for more general rules, authors should consult standard style books.[1]

GENERAL

(1) Spelling. The AMERICAN JOURNAL spells the American way, *e.g. analyze, behavior, counseling, disk, esthetics, gray, questionary, synesthesis,* instead of *analyse, behaviour, counselling, disc, aesthetics, grey questionnaire,* and *synaesthesia,* respectively. The first spelling given in *Webster's New International Dictionary,* 2nd unabridged edition, 1949, is used.

(2)Taboos. Barbarisms, the adjectival use of nouns, and the contraction 'and/or' are taboo. If the avoidance of the adjectival use of nouns leads to involved constructions, the usage is permitted but in that case the nouns should be hyphenated, *e.g. stimulus-object.* The hybrid symbol 'and/or' evidences a poverty of words; it often accompanies barbarisms and the use of nouns as modifiers.

(3) Sex. Animals used as subjects are referred to as *male* or *female;* human beings, as *men* or *women,* or *boys* or *girls.*

PRINTING

(1) Type faces. The JOURNAL uses large capitals (three underlines), small capitals (two underlines), italics (one underline), and italic large capitals (four underlines). Boldface type (wavy underline) is used as occasion demands but it is avoided insofar as possible because it makes the printed page look spotty.

(2) Type-sizes. Three sizes of type are used. The sectional headings (Apparatus, Notes and Discussions, and Book Reviews) are set in 14 point (pt.); the titles of articles in 10 pt. large capitals; the names of authors below the titles in 10 pt. large and small capitals; the introduction, statement of problem, results, discussion, and summary in 10 pt. on 12 pt. matrix (10/12). Historical reviews, descriptions of method and procedure, short articles of six or less pages, and book reviews are set in 8 pt. on 10 pt. matrix (8/10). Footnotes are set on 8 pt. solid, *i.e.* 8 pt. on 8 pt. matrix (8/8).

THE TEXT

(1) Divisions. The text is divided by centrally placed divisions and marginal stubs at the left. The central divisions are set in large and small capitals and if

---

[1] See Publication Manual, *Psychol. Bull.,* 49, 1952, 389-445, that was prepared by the Council of Editors of the American Psychological Association.

numbered, as in reporting of separate experiments, Roman numerals are used. *E.g.*
EXPERIMENT I.

In short articles use marginal stubs throughout to mark the divisions of the paper.
Use stubs also in long articles to indicate subdivisions. The stubs are set in italics
(lower case) and they should be short. They are to indicate subdivisions—not to tell
a story. They should not be assumed in the exposition. The stubs are not numbered
in short articles but in long articles they may be. When numbered, Arabic numerals
(italic), set in parentheses before the subtitle, are used for the first order of sub-
ordination. For the second and third orders of subordination use italic letters of the
alphabet and lower case Roman numerals set in parentheses. For example:

<div align="center">

EXPERIMENT I
</div>

Problem.
Method and procedure.
   (1) Apparatus.
     (a) Accessories.
      (i) Loudspeaker.
      (ii) Microphone.
Results.
      (1) With tones of 8000 cycles and less.
      (2) With 10,000 cycles.
Discussion.

<div align="center">

EXPERIMENT II
(Repeat as much of the above as necessary.)

DISCUSSIONS AND CONCLUSIONS
SUMMARY
</div>

In articles in which there are no center heads, *Discussion, Conclusions,* and *Sum-
mary* are run as marginal stubs.

(2) Titles in text: (a) Articles. Titles of articles and essays, chapters and sections
of books, and unpublished works are enclosed in double quotation marks. The first
letter of the first word is capitalized, all others are set in lower case.

(b) Books. Titles of published books, plays, pamphlets, periodicals, classical
works, and noted poems and operas are set in lower case italics with the principal
words capitalized.

(3) Quotations. All quotations should faithfully duplicate the originals in word-
ing, spelling, and interior punctuation. Any exceptions, such as italicizing certain
words for emphasis, should be explained in an accompanying footnote. When the
quotations are short they are run in the text and set in double quotes; when long
(five or more lines) they are set in a separate paragraph in 8 pt. solid and without
quotes.

For ellipses within quotations use three spaced periods (. . .) and leave a space
before the first period and the last word. If an ellipsis is made between two succes-
sive sentences, use three spaced periods in addition to the period ending the first
sentence.

Interpolations (comments or explanations) within quoted matter should be
enclosed in square brackets, not parentheses. A common interpolation is [sic] used
to invite the reader's attention to the fact that the quotation is accurate.

Quotations from copyrighted books and periodicals legally require permission

from the author or publisher or both. This requirement is ignored for short quotations but for quotations of 300 and more words it should be observed. The task of obtaining the permission is the author's.

Single quotes are used within double quotes. In cases when the quotation is set in 8 pt. solid and the double quotes are omitted, the subquotes are then raised to double quotes.

(4) Punctuation. Do not over-punctuate but use the various punctuation marks as clarity demands.

(a) Commas. In a series of three or more coördinates, use a comma before 'and' and 'or.' The comma should not be used in conjunction with a dash but it should be used after a parenthesis (such as this), if the context requires it.

(b) Dashes. The length of a dash is indicated by the editor. If he writes 'en' above a dash in the text, the compositor will set an 'en' dash, that is a dash as long as the pica 'N.' If he places an arabic 1, 2, or 3 on the dash in the text, the compositor will set a dash that is as long as 1, 2, or 3 pica 'Ms' respectively. The dash below the title and the author's name in the heading of an article is 4 pica 'ems' and is indicated by writing '4' above the dash. One-em dashes should be used to indicate pauses or breaks that are not as complete as those indicated by semicolons.

(c) Exclamation marks. Exclamation marks are rarely used.

(d) Hyphens. Hyphens are indicated by two short parallel vertical lines with short dash midway between them. They ligate two nouns (*e.g.* stimulus-point) into one word to avoid the adjectival use of nouns.

(e) Periods. In addition to closing sentences in text and footnotes, periods end all footnote citations. They are not used after incomplete sentences such as center headings, table captions, and legends to the figures. They do, however, follow marginal stubs.

(f) Parentheses. Parentheses are used to enclose an explanation, authority, definition, reference, translation, or other matter not strictly belonging to the sentence.

(g) Brackets. The use of square brackets is restricted to interpolations within quotations; to editorial comments, corrections, notes, or explanations; and to the numbering of formulas and equations on the righthand margin.

(h) Apostrophe. The apostrophe marks the elision of a syllable (*I've* for *I have*), of the century in dates (*The spirit of '76*), and the possessive case. Elisions rarely appear in scientific exposition. The possessive is indicated by an apostrophe and the addition of the letter *s* for all nouns in the singular number, whether proper or not, and all nouns in the plural ending with any letter other than *s;* as *man's, men's, Charles's, witness's.* The only exceptions are established idioms; *e.g. for conscience' sake, for goodness' sake, for righteousness' sake.* All plural nouns ending in *s* form the possessive simply by the addition of the apostrophe; as, *boys', horses'.* For monosyllabic proper names ending in *s* or another sibilant, add an *s, e.g., Keats's poems, Marx's doctrines.* For proper names of more than one syllable, add only the apostrophe, *e.g. Ebbinghaus' experiments.*

(i) Quotation marks. For quotations within the text, use double quotation marks, then, for quotations within quotations, single marks. Double marks are reserved for actual quotations; they are never used for emphasis nor to designate a special use of a word or phrase—single quotes, or italics, or capitalization are used for those purposes.

Quotation marks are set outside of periods and commas, but inside of colons, semi-

colons, and question and exclamation marks unless the punctuation is actually part of the matter quoted, in which case place the quotation marks outside.

(j) Accents, diereses, and diacritical marks. In quoting foreign articles, duplicate faithfully all accents and diacritical marks. In German words use the umlaut except for initial capitals where *e* is used to indicate it, as *Ueber* instead of *Über*. The dieresis is always used in words with doubled vowels; *as zoölogy, coöperate, reëxamine;* and with *naïve.*

(5) Numerals. Cardinal numbers below 10 are usually spelled out except when they are used before abbreviations; as *two rooms, nine stimulus-objects.* When used before abbreviations, they, in common with all numbers, are printed in Arabic; as *4 cm., 7 ft.* The numbers 10 and above are set in Arabic except at the beginning of a sentence, *e.g. Ninety Ss were used in this study.*

Exceptions to the rule of spelling out 'nine' and numbers below occur when numerals 10 and above occur in the same or contiguous groups of sentences; *e.g. Our Os, 15 in number (7 women and 8 men), were highly trained.*

Ordinal and nominal numbers are usually spelled out. Exceptions occur in the Book Review Section where different editions are indicated, as, *2nd ed., 3rd rev. ed.,* and in the text where points or results are listed in sequence when arabic numbers within parentheses may be used, as (1), (2), (3), etc., in place of *First, Secondly, Thirdly,* etc.

Dates and page-numbers are not spelled out. For dates use month, day, and year, as *October 20, 1954.* References to centuries are spelled out, as *eighteenth century, twentieth century.*

(6) Capitalization: (a) Titles. In English titles of publications, capitalize the first and all principal words; do not capitalize articles, prepositions, and conjunctions. For foreign titles, follow the foreign style.

(b) Abbreviation of psychological terms. Standard abbreviations of psychological terms and others that an author wishes to make to avoid frequent repetition of phrases are capitalized and run in italics without periods. *E.g.* (standard) $O, S, E, SD, CFF, RL, DL, UDL, LDL, PSE;$ (author's) AL (adaptation-level), TOE (time-order effect), CAL (comparative adaptation-level), HMD (half-meridional difference). For the author's abbreviations, see this JOURNAL, 66, 1953, 630 ff., and 67, 1954, 327 ff.

(c) For emphasis. Capitalize such words as *experiments, conditions, trials,* etc., for emphasis—to make them stand out from the text—when they are used in connection with specific experiments, conditions, or trials; as, *Experiment II, Condition 1, Trial 3.* (For examples see this JOURNAL, 67, 1954, 264 f.)

(7) Italics: (a) For emphasis. Use italics sparingly for emphasis that they may be used effectively when occasion warrants them. They lose their significance when frequently employed.

(b) Foreign words. Single foreign words or short phrases are italicized. Foreign sentences are not; they are set in roman double quotes.

(c) Books and periodicals. Titles of books are set in lower case italics with principal words capitalized; titles of periodicals are abbreviated in accordance with the principles adopted by the editors of *A World List of Scientific Periodicals: 1900-1905* and set in italics.[2]

---

[2] For a list of periodicals and their abbreviations, see *Psychol. Abstr.,* 21, 1953, 835-840; for a list of abbreviations, see A publication manual of the American Psychological Association. *Psychol. Bull.,* 49, 1952, 436-438.

(*d*) *Botanical and zoölogical names.* The names of botanical and zoölogical classes, families, genera, and species are italicized. The first letter of the class, family, and genus is capitalized; of the species it is not.

(*e*) *Formulae.* Letters used in formulae and equations are set in italics, upper or lower case as the occasion demands.

(*f*) *Signs.* If the letters used in plates or figures as signs are printed in italics, italics should be used in the text when reference is made to them.

DOCUMENTATION

(*1*) *Footnotes.* Terminal bibliographies are not run in the JOURNAL. All references and digressions (explanations and discussions not of immediate concern to the points being developed in the text) are given in consecutively numbered footnotes.

(*a*) *The logic of footnotes.* Footnotes save space as they are set solid whereas terminal bibliographies are usually leaded. Moreover, only items actually referred to in the text are given in footnotes whereas terminal bibliographies are often padded, *i.e.* items not referred to in the text are frequently included.

References are of greater aid and convenience to the reader if they are placed on the pages of their citations. It is much easier to cast one's eyes to the bottom of a page when a citation is met in the text than to turn to a terminal bibliography.

Footnotes permit of specific citations, *i.e.* the actual pages are given upon which the points referred to in the text appear. Terminal bibliographies give inclusive pages. Reference to an article given in a terminal bibliography is consequently 'blind,' *i.e.* the reader must search through the entire article or book to find the point or points at issue. Such references are of little help to him.

To avoid blind references, journals running terminal bibliographies sometimes number the items in the bibliography and give the number and specific pages in parentheses in the text; as (9, pp. 45-47). If the item-number is printed in boldface, as in some journals, the page appears spotty and is a poor example of the art of printing. Such references, moreover, catch the reader's attention and interrupt the flow of eye movements to a far greater extent than footnote references which are given in small superscripts at natural pauses, *i.e.* at punctuation marks.

(*b*) *References in the text.* Starting with 1, footnote references in the text are numbered consecutively in superscript Arabic figures. Insofar as possible, these figures are placed at punctuation marks, preferably at periods. When clarity (which takes precedence over every other rule) demands it, footnote references may be placed elsewhere.

(*c*) *Footnotes concerned with books and articles.* When the footnotes refer to published books or articles, the following information should be given and in the order indicated below.

(*i*) *Name of author.* The author's surname preceded by his initials (or first name if he has only one given name) is run first. Initials or given name are used only the first time an author is cited unless other authors of the same surname are cited. In subsequent citations to the work of a given author, drop the initials or first name—give only the surname. For example:

[1] E. B. Titchener, Structural and functional psychology, *Phil. Rev.*, 8, 1899, 290-299.

[6] Titchener, The psychological concept of clearness, *Psychol. Rev.*, 24, 1917, 43-61.

Exceptions are made to this rule, however, when subsequent citations are to articles

of plural authorship. If the initials of any of the co-authors of a joint article are given, and they should be if any of the co-authors is cited for the first time, then give the initials for all. For example:

[2] T. A. Ryan, Interrelations of sensory systems in perception, *Psychol. Bull.*, 36, 1940, 659-698.
[4] T. A. Ryan, C. L. Cottrell, and M. E. Bitterman, Muscular tension as an index of visual efficiency, *Illum. Engng.*, 43, 1948, 1074-1081.

If, however, the same or another joint article or book by the same co-authors is cited, then drop the initials of all the authors, as:

[5] Ryan, Cottrell, and Bitterman, *op. cit.*, *Illum. Engng.*, 43, 1948, 1080.
[9] Ryan, Cottrell, and Bitterman, Relation of critical fusion frequency to fatigue in reasoning, *Illum. Engng.*, 48, 1953, 385-391.

(ii) Titles of books. If a book is cited, then continue, after giving the author's initials and surname, with the title. This is set in italics with the first and principal words capitalized. The title is followed by the edition, if other than the first; then the volume-number in Arabic numeral if more than one; the year of publication; and the specific pages. All of these items are separated by commas, for example:

[3] Madison Bentley, *The Field of Psychology*, 1924, 321-323.
[4] E. G. Boring, *A History of Experimental Psychology*, 2nd ed., 1950, 21-24.
[5] William James, *The Principles of Psychology*, 2, 1890, 245.

(iii) Titles of articles. If the article cited is published in a periodical, give, after the author's initials and surname, the title in lower case with the first word capitalized; then the name of the periodical in italics and abbreviated in accordance with the usage of the *World List of Scientific Periodicals;*[3] the volume number in Arabic numerals; the year; the serial number of the publication, if there is one, set in parentheses; and the specific page or page numbers; all items being separated by commas. For example:

[2] M. F. Washburn, The function of incipient motor processes, *Psychol. Rev.*, 21, 1914, 376-390.
[5] Madison Bentley, The psychological antecedents of phrenology, *Psychol. Monogr.*, 21, 1916 (No. 92), 102-105.

(iv) Titles of chapters in a book. If a chapter in a book is cited, give after the author's initials and surname; the title of the chapter in lower case with the first word capitalized; the initials and surname of the author of the book, preceded by 'in' and followed by 'ed.' set in parentheses; the title of the book in italics with principal words capitalized; the edition, if other than the first; the volume number, if more than one; the year; and the specific pages. For example:

[2] Leonard Carmichael, Ontogenetic development, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 304-329.

(v) References to this JOURNAL. If the citation is to an article published in the AMERICAN JOURNAL, use the phrase, 'this JOURNAL' in place of the standard abbreviation. The second word is set in capitals and small capitals; as:

[1] E. B. Newman, The patterns of vowels and consonants in various languages, this JOURNAL, 64, 1951, 369-379.

---

[3] See footnote 2.

(d) Digressions. Digressions, explanations, and discussions, that are not germane to the text, are placed in footnotes. When references are made in them to books or articles, they are given in parentheses in the form shown below.

> ³ The hearing loss was surprisingly less than that obtained by a similar method of impairment. In the earlier study the loss was 65 db. at all frequency-levels and the *Ss.* could not hear their footsteps nor understand speech unless shouted (Supa, Cotzin, Dallenbach, "Facial Vision": The perception of obstacles by the blind, this JOURNAL, 57, 1944, 169.

(e) Repetition of citations. At the first and subsequent repetition of a reference, the abbreviation *'op. cit.'* is substituted for the title. If only one article by a given author has been cited, *op. cit.* is followed by the page numbers. For example:

> ³ S. W. Fernberger, Observations on taking peyote (Anhalonium Lewini), this JOURNAL 34, 1923, 267-270.
> ⁴ Fernberger, *op. cit.,* 269.

If, however, two or more articles by the same author have been cited, the second reference to one of them is designated by *'op. cit.'* and, in the case of books, by the year of publication of the particular book referred to and the specific page numbers. for example:

> ³ Madison Bentley, *The Field of Psychology: A Survey of Experience, Individual, Social, and Genetic,* 1924, 103-105.
> ⁵ Bentley, *The New Field of Psychology: The Psychological Functions and Their Government,* 1934, 187-190.
> ⁷ Bentley, *op. cit.,* 1924, 278-284.

In the case of multiple citations of articles by the same author, the second reference to one of them is designated by *'op. cit.,'* the abbreviated name in italics of the periodical of publication, volume, year, and specific page numbers. For example:

> ¹ K. M. Dallenbach, The temperature spots and end-organs, this JOURNAL, 39, 1927, 402-427.
> ¹¹ Dallenbach, The place of theory in science, *Psychol. Rev.,* 60, 1953, 33-39.
> ¹² Dallenbach, *op. cit.,* this JOURNAL, 39, 1927, 425.
> ¹⁶ Dallenbach, *op. cit., Psychol. Rev.,* 60, 1953, 35.

These rules apply only to citations in short articles and to articles with few footnotes. For the convenience of the reader they may be broken and the complete citation be repeated. Repetition may be desirable in long articles and in articles with numerous footnotes. The reader should be kept in mind; his ease and convenience are considerations that should not be overlooked. If he has to turn back many pages to find the first citation of a given reference, or if he has to search through numerous references to find it, the repetition of the title is advisable.

ABBREVIATIONS AND REFERENCE WORDS

Abbreviations and reference words may be divided into two classes: those reserved for the editor, and the standard ones that may be used by the author.

(1) Editor's abbreviations. The following reference words and abbreviations of Latin words should be left to the use of the editor.

cf., abbreviation for *confer,* meaning 'compare.' Do not use *cf.* when 'see' is intended. It is italicized.

ibid., for *ibidem,* meaning 'in the same place,' *i.e.* in the reference immediately preceding. It is italicized and

should always be followed by specific page numbers.

idem, this is a reference word, not an abbreviation. It means 'the same,' *i.e.* the same person as cited in the reference immediately preceding. It is italicized. Though this reference word was for a time frequently used in the JOURNAL, it is now, under the rule described above (Section *1, c, i),* rarely used. The author's name is instead repeated. This does not consume much more space and is clearer as the citation referred to may be on a preceding page and the reader, to discover it, may have to turn back. The repetition of the name avoids this inconvenience.

loc. cit., (plural *locc. citt.*) for *loco citato,* meaning 'in the same place

cited,' *i.e.* in the same passage referred to in a recent note. It is italicized but is not followed by a page number. The page number or numbers given with the earlier reference are implied. If the same article but different pages are to be indicated, then *ibid.* or *op. cit.* should be used, depending upon which of the two is appropriate.

op. cit. (plural *opp. citt.*) for *opus citatum,* meaning work cited.

p., (plural pp.) for 'page.' Set in roman, not italicized.

q.v., for *quod vide,* meaning 'which see.' Rarely used in the JOURNAL.

vol., (plural vols.) for 'volume.' Set in roman and omitted when preceded by the name of the publication and followed by the year and inclusive pages.

The use of these tools of reference is restricted to the editor because he is the best judge as to which one of them, if any, should be used. Authors should not use them; they should complete their reference as it is much easier for an editor to cross out and interpolate the correct abbreviation than it is to cross out the incorrect abbreviation and substitute the correct one, or to repeat the reference as he may deem advisable.

(2) Standard abbreviations. Among the abbreviations frequently used in the text and footnotes are the following:

A.C., (alternating current), small capitals with periods but without spacing. Used only when preceded by a specified voltage; as 25 v. A.C.

A.D., (*anno Domini*), small capitals with periods but without spacing. Precedes numerals. Used only in contradistinction to B.C.

AD, (average deviation), italic capitals, without spacing.

A.M., (*ante meridian*), small capitals with periods but without spacing. Follow numerals denoting the hour.

anon., (anonymous).

AQ, (achievement quotient), italic capitals, without periods and spacing.

b., (born).

B.C., (before Christ), small capitals with periods but without spacing. Follows the numerals.

CA, (chronological age), italic capitals, without spacing.

ca., (*circa,* about). Precedes approximate date.

CFF, (critical fusion frequency), italic capitals, without spacing.

cycle, Spell out except when preceded by

a numeral, then use sign ($\sim$).

d., (died).

db., (decibel). Spell out except when preceded by a numeral.

D.C., (direct current), small capitals with periods but without spacing. Used only when preceded by a specific voltage.

Dr., (doctor). Always use as a title before a name.

d.v., (double vibration), lower case with periods but without spacing. Used only when preceded by a numeral. Cycle sign is preferable.

E, Es, (experimenter, experimenters), italic capital. Possessive, *E*'s, *Es*'.

ed., eds., (editor, editors).

ed., (edition), preceded by the number of the edition other than the first; as, 2nd ed., 3rd ed.

e.g., (*exempli gratia,* for example), lower case italics without spacing and not followed by a comma.

etc. (et cetera, and so forth), this abbreviation should be used sparingly. It is of little help to the reader except as it denotes an orderly continuum.

f., ff., (and the following page or

pages). Preceded by a page number and a space. A single 'f.' means one page; double 'ff.' two or three pages. If more pages are intended, give inclusive pages.

Fig., Figs. (figure, figures). Always capitalized and followed by an Arabic number or numbers. Spell out when not.

g., (general factor).

GSR, (galvanic skin response), italic capitals, without spacing.

gm., (gram).

h, (index of precision).

i.e., (id est, that is), lower case italics, without spacing and not followed by a comma.

IQ, (intelligence quotient), italic capitals, without spacing.

Jr., (junior), follows name and is part of it. It takes the possessive, as "John Smith, Jr.'s study."

M, (arithmetic mean).

m., (meter).

MA, (mental age).

Messrs., (misters).

min., (minute).

mm., (millimeter).

Mr., (mister); never spelled out except with humorous intent.

Mrs., (mistress).

m.sec., (millisecond).

MS, MSS, (manuscript, manuscripts). Capitals without periods and spacing.

N, (number), mathematical symbol referring to number of cases.

N.B., (nota bene, mark well), capitals with periods but without spacing.

no., nos., (number, numbers).

O, Os, (observer, observers), italic capitals. Possessive, O's, Os'. For difference between O and S, see Madison

Bentley, 'Observer' and 'subject,' this JOURNAL, 41, 1929, 682 f.

ohms, Spell out except when preceded by a numeral, then use capital Greek omega (Ω).

PE, (probable error), italic capitals without periods and spacing.

percentage, Spell out except when preceded by a numeral, then use percentage sign (%).

P.M. (post meridian), small capitals with periods but without spacing. Follows numerals denoting the hour.

Professor, Always spell out.

PSE, (point of subjective equality), italic capitals without periods and spacing.

R, (Reiz = 'stimulus' in psychophysics, italic capital).

R, (response in behavorial psychology, roman capital).

rev., (revised), used before 'ed.'; as '2nd rev. ed.'

RL, (stimulus limen), italic capitals without periods and spacing.

r.p.m., (revolutions per minute).

r.p.sec., (revolutions per second).

S, Ss, (subject, subjects), italic capitals. Possessive, S's, Ss'. For difference between S and O, see Bentley, op. cit., this JOURNAL, 41, 1929, 682 f.

S, (stimulus in behavioral psychology, roman capital).

SD (standard deviation), italic capitals, without spacing.

sec. (second).

sic (thus, so), between brackets in quoted matter when used as an interpolation; otherwise within parentheses.

TR (terminal stimulus).

trans. (translation).

viz. (videlicet, namely), roman, with period but not followed by a comma.

Abbreviations for times, weights, and other measures are the same for the plural as for the singular. Context determines which is intended; as 1 sec., 6 sec.; 1 gm., 10 gm.; 1 lb., 15 lb.; 1 ft. 7 ft.; 1 mm., 35 mm.

TABLES AND FIGURES

(1) Tables. The tables within every article are numbered consecutively in capital Roman numerals; as Table I, Table V. When the tables are referred to in the text by number, the first letter is capitalized, otherwise it is not. The word 'Table' and its number are set in capitals and centered on the page above the tabular matter. This is followed by an explanatory caption, set in capitals and small capitals, and centered upon the page. The caption should be brief but sufficiently detailed to permit the reader to discover what the table purports to show. The caption may have interior punctuation but it should not end with a period. Additional explanatory

material may be given in parentheses centered under the caption. This is set in lower case. Footnotes to the table, if required, should be referred to by the following printer's signs and in the order listed: asterisk (*); dagger (†); double dagger (‡); and sectional sign (§). For example:

* Significant at the 1% level.
† Significant at the 5% level.
‡ A paper-and-pencil version of the test used here.
§ Performed while standing and not on treadmill.

(2) Figures. The figures (line and half-tone blocks) within every article are numbered consecutively in Arabic numerals. The word 'Figure,' both in the legend and when referred to with its number in the text, is abbreviated; as Fig. 1, Fig. 2, Fig. 5. In the text, the first letter of the abbreviation is capitalized; under the figure, the abbreviation and explanatory legend are set in small capitals with principal words in large capitals. The legend does not end with a period; it is not a sentence. If further explanatory material is needed, give it in lower case, centered under the legend.

The figures should be drawn in India ink, with the size of the JOURNAL page in mind. It is 25 × 42 picas (approximately 4 × 7 in.). They should be drawn by a draftsman and lettered so large that they will easily be read after the necessary reduction. Amateurish drawings should not be submitted. Since reproductions can usually best be made from original drawings, these are preferred, though glossy prints of satisfactory drawings are acceptable.

The University of Texas   KARL M. DALLENBACH

BOOK REVIEWS

Gesetze de Sehens. By Wolfgang Metzger. Second edition. Frankfurt am Main, Waldemar Kramer, 1953. Pp. xii, 470.

Metzger's book, already nearly three years old in the present edition, appeared in other dress still earlier. Most of it was originally presented in the form of individual articles in a magazine published by a museum of natural history. In 1936, a small but uniquely valuable book was made from one series of those papers. In 1953, the present, much expanded version appeared. The mode of presentation is not easy to categorize. The work is neither textbook nor handbook; it is written more for the scientifically curious than for the scientifically sophisticated; its tone usually is colloquial and practical rather than academic and theoretical. Nevertheless, it is in certain ways the best available book on visual perception. It presents a wealth of material on most of the major problems in that area; it treats the available facts fully; and it provides information that is rarely brought together in one place. Though not written for classroom use, it teaches, and teaches systematically. It develops the capacity to ask questions. Readers who have been taking their visual world for granted will never do so again, and even some who have studied visual problems for a long time may find new kinds of question arising as they read.

The most striking feature of this book is its profusion of illustration. There is hardly a page among the 470 that does not contain some sort of photograph or drawing. Wonderfully useful photographs abound, as they did in the first edition: photographs of paintings that illustrate the effects of chiaroscuro and of perspective, photographs of protective coloration and concealment in the animal kingdom, photographs of laboratory-demonstrations in the field of form or color, and many others. The volume is still more lavish in its use of line-drawings to clarify principles of stereoscopy, of grouping, and other psychological phenomena, as well as to portray experimental ararngements in detail or to schematize hypothetical situations.

Most remarkable is an unusual array of demonstrations that are made possible by two-color printing. Binocular vision becomes a subject that is demonstrated as well as discussed when the reader makes use of the red and green filters provided with the book. Equally effective and even more unusual is the use of the same principle to quicken the topic of perceived movement. The printing of one part in red and another in green has enabled the reader with a pair of red and green filters to alternate the exposure of various patterns that illustrate almost all the major problems and effects. Where most published accounts are useful only to readers who have already seen demonstrations elsewhere, Metzger has made it possible for the beginner to follow work like that of Ternus or Schiller with full understanding.

While one's first impression of this book is that it is a triumph of the bookmaker's art, a few complaints should be registered in the hope that they may lead to improvements in later editions. The printing is not up to that of the first edition. Several potentially useful photographs are spoiled by poor reproduction. The construction of the binding is shoddy; this book will not be able to stand up to the

amount of use that it will attract. Most annoying of all is the complete lack of an index, along with a corollary confusion and lack of system in textual citations. Some of the latter can be found in a list of references at the end of the chapter; others are found only under illustrations; still others appear in footnotes. Some of these matters could be cleared up with very little effort; doing so would add to the value of the book.

On grounds of content, this book could be highly praised for what it includes or roundly criticized for what it leaves out. I am far more inclined to the former course. Metzger presents a great deal of material here that has not previously been available in book-form. Most of the important German work of the 1930s is now available in one place, fully and clearly treated in the context of a theoretical survey of vision. That is in itself a major contribution. In addition, there are special things: notably the fine accounts of concealment in animals as illuminated by psychological principles. Experiments and demonstrations that have been all but forgotten by American students are again brought to life—Madlung's observations on spatial organization in touch, Musatti's on the stereo-kinetic phenomenon, Krolik's on learned frames of reference for perceived movement, and others. The first hundred pages of the book would alone be worth the purchase-price. In them the author treats the basic problems of visual patterning and organization, making the reader really examine his visual world and then showing how difficult it is to accept any of the usual interpretations of the nature of seeing; how objects are 'created' by the visual system; how they may also be made to disappear; how the perceived world seems to have a structure that is dependent neither on what 'it really is' nor on the way we have learned it to be. The sections on size and distance are almost as good. Psychologists brought up on traditional theories about binocular vision and about perspective will find illustrations here that they do not find in ordinary accounts, illustrations that enrich the application of psychological thinking to the arts and to practical living. The chapters on movements, too, are unusual. The material that they incorporate is not new, but so many different aspects of visual movement have been brought together in one place that we get the impact of an original treatment of the problem.

Most of these same things have been done before by Wertheimer or Koffka or Köhler, but they have never been done so extensively or with such lavish illustration. The very completeness of the exposition adds enormously to its effectiveness. Yet, in another sense, it is not really complete. Almost no American or English work is included and almost nothing from post-War researches anywhere. All of our recent controversies about the influence of motive, set, attitude, and special forms of past experience are omitted. Many relevant discoveries and hypotheses that have become familiar to workers in the English-speaking countries are missing completely.

There are two ways of looking at such omissions. Since this is not meant to be a reference book or encyclopedia, one may readily argue that we need only those facts sufficient to illustrate basic principles; detailed controversy or elucidation of technical points would be out of place, and the older literature is sufficient. On the other hand, Metzger is presenting a coherent theoretical position, and it might be argued that he makes his case spuriously tight by ignoring certain kinds of facts. I am willing to accept the first viewpoint, though with certain reservations.

Metzger presents what we have come to know as the Gestalt point of view. He proposes, following Wertheimer and others, that visual perception has laws of its own, laws of organization that cannot be derived from a molecular analysis of stimulation to which principles of association, attention, and the like are later added. He makes of perception a branch of general psychology, minimizing those influences that permit different people to see different things in the same stimulus-constellation. He has prepared an impressive array of facts to support his case, bringing together all the well-known ammunition plus a good deal more, and adding to the effectiveness of it all with his style of reasonableness and common-sense. In certain areas, it is hard to quarrel with this approach; the phenomena of grouping, of figure-ground organization, and one or two others seem to lend themselves to no other point of view without considerable strain. In other fields, conclusions may have to be drawn more cautiously. For example, Metzger treats the well-known distorted room of Ames as a case of "good Gestalt" with no bow whatsoever to the view that the phenomenon demonstrates the influence of past experience (and, incidentally, without citing Ames or his collaborators). Readers may well come to the conclusion that those who view perception as idiosyncratic, a form of behavior primarily determined by individual life history and contemporary motivation, are old-fashioned people who do not know the facts, and that we scientists have long since disposed of them. I wish it were true! On the balance of the available evidence, I think Metzger's position is the more nearly correct one, but I am sorry that the opposition is not given a better opportunity to state its case.

Nevertheless, Metzger's method has real virtues. By ignoring evidence that he considers irrelevant or incompetent, by not troubling the reader with issues that he feels are trivial or too technical or still unsettled, he manages to present a unified, consistent treatment that epitomizes an important theoretical position. No previous account has provided so thorough a display of the application of Gestalt-theory to visual investigation. Perhaps it is enough to provide such an exhibit for the more critical to survey or to shoot at; at least, one now knows where to send students or critics for an adequate elementary account of the Gestalt approach to almost any problem in visual perception.

The usual biases and limitations of the Gestalt viewpoint are, of course, present. *First*, most of the facts offered are phenomenological and qualitative. There is almost no attempt to provide independent criteria of what is being perceived, and there is very little of the experimental rigor that most American psychologists like. There is little or no real measurement of the extent of the various visual effects or of the stimulus-changes that are supposed to be correlated with those effects. *Second*, numerous perfectly good psychological issues in the area of vision are completely ignored, presumably because they do not lend themselves readily to an organizational interpretation—*e.g.* visual acuity, color-sensations and mixture, dark-adaptation, eye-movements, and others—are barely mentioned. The theory selects the facts as much as *vice versa*. *Third*, questions about the structure and function of receptors are not raised with any frequency or in any detail. Even the hypotheses about cortical function that have developed in latter-day Gestalt thinking are omitted.

Within these not-too-confining limits, it is hard to see how Metzger could have

done a much better job. His chapters are generally well-organized and clear. They are written in a brisk declarative style with frequent references to pictorial material and to everyday life. They provide citations for a sufficiently large proportion of the material to enable a tyro to begin to discover the literature of vision and to enable an expert to locate most of the original research. The book is provocative, and it is fun; it is thorough within its own self-imposed limitations, and it is sound. I wish it were available to every American student of perception. The simplest way to express my over-all evaluation, perhaps, is to say that it must be translated. Unfortunately, the advent of World War II prevented completion of a project to publish a translation of the much smaller and less ambitious first edition. The present volume has all the unique merits of the older one and many more. No book presently available in English performs the same service.

Swarthmore College   W. C. H. PRENTICE

Mathematical Thinking in the Social Sciences. Edited by PAUL F. LAZARSFELD. Glencoe, The Illinois Free Press, 1954. Pp. 444.

Lazarsfeld, like many another scientist, is convinced that the times require "the development of a creative and useful science of human behavior and social relations," and he is convinced that more enthusiastic and sophisticated applications of mathematical tools will speed this development. The testing of the latter proposition requires research by social scientists trained in mathematical methods or by social scientists and mathematicians in collaboration, but it requires also, in Lazarsfeld's view, examination by social scientists in general of examples of the ways in which "the structure of behavioral science thinking and the structure of various mathematical methods fit each other." It is in this examination that he sees the contribution of the present volume.

The core of the collection apparently derives from a series of invited addresses sponsored by the Department of Sociology at Columbia University. Nicholas Rashevsky contributed three lectures to this series, two of which appear in the present collection under the title "Two Models: Imitative Behavior and Distribution of Status." Jacob Marschak apparently contributed three lectures, all of which appear under the title, "Probability in the Social Sciences." Louis Guttman and Lazarsfeld both contributed, and in this volume are two papers by Guttman ("The Principal Components of Scalable Attitudes," and "A New Approach to Factor Analysis: The Radex") and one by Lazarsfeld, "A Conceptual Introduction to Latent Structure Analysis." In preparing these lectures for publication, Lazarsfeld solicited several others to round out the volume: a paper by J. S. Coleman designed to clarify for the non-mathematical social scientist the contributions of Rashevsky (but not solely his contributions to this volume), a paper by T. W. Anderson on a Markov-process model of attitudinal change, and a paper by H. A. Simon on "Some Strategic Considerations in the Construction of Social Science Models."

Although Lazarsfeld attempts in his introduction to integrate the several papers, they remained, for this reader, essentially independent as between authors, except, of course, for the Rashevsky-Coleman series. In view of the very wide range of content which term "social science" subsumes, it would probably be impossible to provide truly integrated and representative case-studies of the entire content-area.

This comment should not be interpreted to mean, however, that the problems of the social sciences have here received "representative" treatment in the Brunswikian sense. Of the eight contributions listed in the table of contents, four (Anderson's, both of Guttman's, and Lazarsfeld's) must be considered primarily psychological, and three of these are intimately concerned with attitudes. Two of the eight contributions (Marschak's and Simon's) are primarily economic—though both seek, with some success, to transcend economics. The remaining two (Rashevsky's and Coleman's) are at a boundary between psychology, sociology, and mathematics, if this boundary may meaningfully be said to exist.

I found the papers by Marschak on probability to be the most exciting, and I heartily commend them to all social (and biological) scientists who have any residual curiosity as to the rôle of probability-concepts in their areas. Indeed, I recommend them to every one who is either (*a*) satisfied with the rôle of classical probability theory in his science, or (*b*) satisfied that probability theory has no place in that science—if there still be any such persons. This recommendation is based, not as much on the informative value of the papers (which is considerable) but on their leavening potential. It is not to be expected that the notions of utility and subjective probability here discussed will be uncritically accepted by psychologists, but the very careful consideration given these notions by Marschak (and, at a different level, by L. J. Savage his colleague on the Committee on Statistics at the University of Chicago, in *Foundations of Statistics*) will stimulate the psychologist to new ways of thinking about old problems and will raise new problems for him. Marschak's papers impressed me once again with the fact—and the pity of it— that the psychologist may all too easily ignore that work of the economist which (especially where the economist uses mathematical methods available to the psychologist) can be translated into application to psychological problems. The conviction that mathematical methods facilitate translation across disciplinary borders was one of the editor's motivations for this volume, but Marschak's reference to the work of Koopmans, of Reiersøl and Koopmans, and of Neyman, on the identifiability of the structural parameters of mathematical models in economics—work with which psychologists seem not to be familiar—points up again the well-known and unfortunate aspects of departmentalization.

Guttman's paper on the principal components of scalable attitudes extends his earlier discussion in *Measurement and Prediction*. Here he briefly reviews that discussion and considers the psychological significance of the several solutions of the set of equations which the concept of a perfect scale implies. He discusses several empirical investigations of the validity of the psychological implications inferred from his mathematical model. The final few sections of this paper—a kind of autobiography of a psychological theory (a premature autobiography, to be sure)—is particularly revealing of the rôle played by mathematics in psychological theorizing. The paper on the Radex (90 pp.) seeks to order, under a single formulation, many points which have formerly divided factor analysts. Only the expert in this field can say how successful this effort has been, and the paper deserves review by such an expert at greater length than would be appropriate here.

Now it is time to comment on the impression with which this volume left me as to the prospective usefulness of mathematical models in the social sciences. In reading those papers in which mathematical models are presented (*e.g.* Anderson's

paper on "Probability Models for Analyzing Time Changes in Attitude," and Rashevsky's papers), I could not escape the reaction—and old friend—that while the models are stimulating, exciting, and fun, the mathematics applied is capable, at the moment, of dealing only with the psychology (sociology, economics) of those ultra-simplified robots which man can now construct; that the real persons and societies which the social scientist seeks to understand are characterized by more parameters mutually influencing each other in ways much more complex that those allowed by the models. Thus the question again arises (and I oversimplify somewhat) as to whether the approach by way of quantitative symbolic logic (of which mathematics in its various branches provides the most familiar formalized examples) which has served the physical sciences so well is capable of similar service to the social sciences. This is a very old question. Lazarsfeld, Anderson, Rashevsky, Marschak, Guttman, and Simon clearly believe that this approach will serve the social scientist well, and so do numerous other social scientists, to many of whom only the simplest of the mathematical tools are available. The answer to the question as to the direction this service must take seems to lie in the relation between the formalization of symbolic logic (mathematics) available on the one hand and the problems of the science on the other. Modern formal mathematics has not grown up independently of the sciences it has served, and it is doubtful if there exists today a competent mathematician who cannot discuss some aspect of physical theory with a specialist in one of the physical sciences. The need for 'new mathematics' to deal with problems of social science—a need which many have seen—is, then, a need for the formalization of a quantitative symbolic logic more congruent with the (more complex?) problems of social science with which it will have to deal. From this point of view, an early symbiotic association of mathematician and social scientist, or of mathematics and social science, is a goal the desirability of which the social scientist should be ever ready to preach to his friends, mathematical and socio-psychological.

Perhaps, however, the question should be stated differently: Can the social sciences be expected to proceed at a tolerable rate if its agents restrict themselves to what Lazarsfeld at one point incidentally refers to as "discursive" methods? Many, including this reviewer, are convinced—at least on Mondays, Wednesdays, and Fridays—that they cannot. This is not, however, to say that the social sciences are at the place where it would be desirable (even were it possible) suddenly to give every worker in the field complete familiarity with the calculus, theory of functions, differential equations, finite differences, Galois theory, matrix algebra, topology, and so on. The momentary potential of mathematical methods for a science likely depends upon the current state of the science. It may be that in every field a certain amount of "discursive" thinking must be done before that explicit recognition and precise formulation of postulates and primitive terms which the mathematical approach requires shall be possible—and this is a matter quite independent of the "congruence" of formalized mathematics with the problems of the science.

Yet it is those who are prepared who will exploit that time when it arrives in any branch of science, and a careful perusal of the volume at hand is likely to lead every psychologist to consider again—and perhaps somewhat more carefully than he has done before—the question of what mathematical training we ought to require of ourselves and of our neophytes.

University of California

RHEEM F. JARRETT

La Perception de la Causalité. By A. MICHOTTE. Second edition. Publications Universitaires de Louvain, 1954. Pp. vii, 306.

Many American psychologists, reared to our tradition of tough-mindedness, may glance at the title of this book, translate it much too literally, wince at its apparent philosophical overtones, and promptly consign it to the file-and-forget tray. That would be a genuine misfortune. The author points out in the preface to the second edition (which is merely a reprinting, with typographical improvements, of the first) that the title has given rise to some misapprehensions, in that it seems to impart an aura of metaphysical or epistemological speculation which even a cursory reading would surely dispel. This book is in effect a long monograph, reporting in detail the results of an integrated sequence of parametric studies in a relatively neglected area of visual perception—the whole class of phenomena characterized by the apparent mechanical action of one visual object upon another. As such, it is well worth the serious attention of students of perception. Michotte's canons of evidence are those of the scientist. The experiments are ingeniously devised and carefully performed. It is only the problem which at first seems exotic.

Michotte develops at great length and in impressive detail the results of a dual enterprise, experimental and theoretical. The experimental task was to produce, under laboratory conditions, visual displays which would give rise to the phenomenal appearance of causal action, and to investigate the influence of relevant stimulus-variables. In this task he has succeeded admirably. With some exceptions, his theoretical statements take the form of first-order empirical laws quite closely tied to the data. Michotte asserts that his conception of the structural laws related to the phenomenon of apparent causation has permitted accurate prediction of the appearance or nonappearance of this phenomenon in novel, presumably critical, experimental situations. Unfortunately, none of this new work has found its way into the second edition of the book, although some scattered reports of relevant research emanating from the Louvain laboratory have appeared in the literature of the last five years.

Two paradigmatic experiments will convey some notion of the general character of the work. These experiments, or demonstrations, give rise to visual experiences typical of the two generic types of phenomenal causation, the "billiard ball" effect (*l'effet Lancement*), and the "pushcart" effect (*l'effet Entraînement*). In the first, *O* is stationed 1.5 m. from a slotted screen. The slot is 150 mm. long and 5 mm. high. Behind the slot, two 5 mm. square, colored patches (*A* and *B*) are visible on a white ground, one in the center of the slot and the other 40 mm. to the left. As *O* fixates *B*, the center patch, *A* begins to move at the rate of 30 cm./sec. in the direction of *B*. When the leading edge of *A* meets the trailing edge of *B*, *A* stops moving and *B* begins to move off to the right, usually at the same speed as *A* formerly moved. *B* stops moving after having been displaced 20 mm. or so. Almost without exception, hundreds of *O*s have reported that they saw *A* approach and strike *B*, sending *B* off to the right. This visual effect, which Michotte calls *Lancement*, is quite compelling and is not overcome by hundreds of repetitions with the same *O*, nor by prior acquaintance with the technique for presenting the display.

The second demonstration is much like the first, except that *A* does not stop when it reaches *B*, but continues at its original speed. Patch *B* begins moving at the same speed as *A* from the moment of contact, and the two patches move side by side

once both have entered into motion. In this case, which Michotte calls *Entraînement*, *Os* report seeing *B* swept along or pushed forward by *A*. Both in this effect and in *Lancement*, patch *A* is seen as active and moving spontaneously, while *B* is seen as passive and being displaced only by virtue of *A*'s movement. It is this peculiar impression, namely, that the appearance of movement as such seems to be a visual property of the active particle only, which constitutes the defining property of these phenomena as a class.

The spatial displacement of the passive particle is seen as somehow attributable to its participation in the movement of the other. Within fairly well defined spatial limits, it is not seen as moving in its own right. This is true even when, in *Lancement*, the active particle is at rest. In attempting to explicate this oddity, Michotte makes one of his few excursions into rationalistic theorizing. A large portion of his major theoretical exposition is given to an intricate discussion of the logic involved in his account of the observation that the movement of a given object can persist phenomenally, by "extension" to another object, after its physical motion has in fact ceased. The argument here is labored and reminiscent of Scholastic disputation, although Michotte blunts this criticism by suggesting it himself. His defense is that the minute complications of his argument are demanded by the complexity of this hard-to-understand phenomenon. Some readers will question the necessity of describing the phenomenon in such complex terms to begin with.

The bulk of Michotte's book is devoted to reporting the results of nearly a hundred experiments, most of which were performed in order to elucidate one or another aspect of *Lancement* and *Entraînement*. In doing this thorough and painstaking job, the author systematically manipulated almost every factor which might conceivably affect the occurrence or mode of appearance of the phenomena. Michotte's general plan of attack is an interesting one. The primary objective was to study a sufficient number of parameters to permit the precise and detailed specification of the necessary and sufficient conditions for the occurrence of the typical causal impressions. Once identified, these conditions were simplified in a manner which made possible an orderly comparison of the phenomenal impressions produced in the simpler situations with the causal impression itself. The method is analogous to that of comparative anatomy. Appearances homologous to those being elucidated are sought in rudimentary phenomena of the same general class. The resulting over-all impression is of a massive and well built edifice of experimental findings.

Michotte is most explicit in asserting that the causal impression is experienced directly and immediately. It is reported to be neither inference, interpretation, nor meaning superimposed upon the perception of two separate but spatiotemporally related movements. For Michotte, the causal impression is a unitary perceptual datum. His general adherence to the Gestalt point of view is in evidence here, as well as in his insistence on the unimportance of past experience as a determining factor. Evidence for the latter point is adduced from a number of so-called paradoxical cases in which the causal impression arises or is enhanced under conditions which are absurd from the standpoint of elementary mechanics and common sense alike.

Considerable space is devoted to matters other than the experimental and theoretical analysis of *Lancement* and *Entraînement*. The apparatus and techniques employed in presenting the visual displays are discussed at length. There is a chapter relating some attempts, mostly unsuccessful, to evoke an impression of the production of a qualitative change in one object by the movement or qualitative change of

another. Two chapters are devoted to the production of more "lifelike" appearances of phenomenal causation by means of simple stimulus-analogues of archery, javelin throwing, and so on. These observations are nicely tied to the basic effects. A few fascinating demonstrations of the possibility of mimicking simple patterns of animal locomotion with stringently reduced stimuli are described. There are some lengthy speculations concerning the role of kinesthesis and the body-image in the production of causal impressions arising from contact between the human body and objects external to it. Michotte develops in the final chapter his opinions concerning the manner in which his results bear upon the views of Hume, Maine de Biran, and Piaget concerning the origin of the notion of causality. One further variety of causal impression, that of a moving object appearing to leave a trace behind it, somewhat embarrassingly fails to fit nicely the general theoretical scheme into which Michotte has cast the major phenomena of *Lancement* and *Entraînement*. The implications of this seeming failure of the theory are squarely faced, but the discussion is relegated to an Appendix.

It is a refreshing, though all too rare, experience to read the report of an experimental enterprise the publication of which has awaited the completion of a truly exhaustive investigation. Every lead has been picked up and patiently run to earth. The result is a monument to the scientific skill and scholarship of the author.

University of Pennsylvania   JOHN WERNTZ

Handbook of Social Psychology: Vol. I, Theory and Method; Vol. II, Special Fields and Applications. Edited by GARDNER LINDZEY. Cambridge, Addison-Wesley Publishing Company, 1954. Pp. x, 1226.

During the short year since its appearance, Lindzey's *Handbook* has won quick and enthusiastic approval. Well publicized in advance, known to contain chapters by the country's outstanding social psychologists, eagerly snapped up by *Basic Books* as a work to be recommended, its formidable price has not kept it from the bookshelves of libraries, scholars, and graduate students. Seldom has such a weighty and expensive publication been received with so little resistance. Now that it is settled on the bookshelf of every serious social psychologist it bids fair to take command of the field. Is this a good thing or a bad thing for social psychology? Rather than assess the contents of the work in detail, the reviewer examines some of the implications of such a publication.

For those few who have not seen the two volumes, one may safely report that the work is brilliantly conceived, magnificently edited, and beautifully published. No teacher of social psychology can do without it. Volume I begins with the finest digest of the history of social psychology that has yet appeared, revealing Gordon Allport at his scholarly best; continues with a series of compact chapters on contemporary theoretical positions—S-R, cognitive, psychoanalytic, field, rôle—all well buttressed by references; and culminates in a number of chapters presenting the different methodological approaches—controlled experimental, sociometric, extensive, and so forth. Volume II leads into some of the less disciplined, but perhaps more appealing, fields of social psychological investigation. In a group of five chapters on "the individual in a social context," the various aspects of socialization are considered, including such problems as social preception, language, and humor. A

further section surveys the literature on group behavior, ranging from the dynamics of small groups to the psychological properties of nations; and four final chapters deal with the applications of social psychological theory and method to the solution of major social problems. The two volumes constitute an extraordinarily comprehensive presentation of social psychology as it is today.

There are, in all, 30 chapters. Each has been prepared by a distinguished scholar or group of scholars. To run through the roster would be to list the best known senior contributors to American social psychology, plus an excellent sampling of younger people who are beginning to win reputations. One would like to pay compliments to all the authors, for the level of writing is almost uniformly high; but this would make for dull reading. Allport's chapter has already been mentioned. Among the other contributions that have particularly impressed this reviewer are: Lambert's brilliant synthesis of the various S-R approaches, Scheerer's valiant attempt to make the word "cognitive" stand for a theory, the short but admirable summary of quantitative techniques prepared by Mosteller and Bush, Berelson's honest but challenging assessment of content analysis as a method for social psychology, the invitation from Hebb and Thompson to take another look at the animals, the refocussing by Bruner and Tagiuri of the psychology of perception on the problems of perceiving people. More space would permit more compliments; and other readers will have different preferences. The simple fact is that Lindzey has marshalled an extraordinarily good group of authors, who have conscientiously adhered to an intelligent plan, and who under his guidance have produced the best composite work on social psychology that has ever appeared. Psychological scholarship owes a great debt to Gardner Lindzey.

To return to our question, however: Is this what we really need in social psychology? The reviewer admits that he snapped at the *Handbook* with glee. Here is a simple way of disposing of the young graduate student in quest of knowledge about social psychology. Give him 1226 pages of scholarship to ponder over, and he will not come back for months. Then again, it solves the problem of the seminar on Social Psychology. Chapter by chapter it will take care of the seminar for a whole year. Why should we professors take time out to mull through the literature, worry about the problems of the field, and try to find a structure that makes sense? Lindzey has given us a structure, and he has given us enough references to keep busy even the most avid of graduate students.

The reviewer confesses that he views a handbook of social psychology with some alarm. Handbooks are certainly useful, but they can also be dangerous. In a well developed branch of science, a handbook can provide a dull but convenient summary of non-controversial facts. In a science that is still groping for principles and methods, it may prove to be a refuge for the intellectually lazy, and it may even arrest rather than stimulate further development. The very bulk of the work may weaken the impulse of the student and the teacher to consult original sources; and the roster of distinguished authors may create a false impression of solidity. The most solidly established branch of psychology is what we call "experimental psychology," and Stevens has given it an admirably edited handbook. Yet what experimental psychologist is happy about the presentation of his own special field? Most of us are grateful for the volume, because it is convenient, but we shudder at the thought that untold thousands of students in other universities

will accept the Stevens presentation as final. We fear that for a generation to come many of our own cherished problems and theories will be overlooked.

This is bad for experimental psychology. Is it equally bad for social psychology? The reviewer's hunch is that it is dangerous, but in this case not necessarily bad. Lindzey's *Handbook* may freeze a whole generation of social psychologists into the Lindzey pattern—and it is really not a bad pattern. As one reads the chapters, however, one gains the impression that the authors, while occasionally uncomfortable with the topic set for them, are genuinely wrestling with problems. They are not presenting "the facts" of an established science; they are playing with alluring questions. It is unfortunate, for instance, that Lindzey's chapter heading would suggest that there is a "cognitive" theory, but a close reading of Scheerer's chapter makes it amply clear that no such theory exists. There are theories of cognition, and there are cognitive approaches to the theory of social behavior, but cognition is a process and a problem, not a theory. Similarly, the reader of Sarbin's delightful chapter on rôle theory will find, not a dull and competent presentation of an established position, but an invitation by way of the theory of social rôles to further social psychological investigation. Even the methodological chapters are studded with unanswered questions; and the 'applied' chapters, like Haire's on industrial social psychology, keep reminding us of problems that have not yet been solved. For obvious practical reasons Lindzey has had to categorize the field of social psychology, but he and his authors have been signally successful in presenting social psychology as a science of unsolved and alluring problems.

This seems to add up to the conclusion that those who really study Lindzey's *Handbook* will not find that it constricts their thinking. It will, on the contrary, stimulate them to further inquiry. Social psychology is not yet an orderly and compact science. It has problems that are too big for its methods, and methods that are often too precise for its data; hence its challenge. Lindzey is to be complimented for having successfully maintained, throughout his immensely difficult editorial task, the conception of social psychology as a growing and challenging science.

Cornell University   R. B. MacLeod

Emotions and Bodily Changes: A Survey of Literature on Psychosomatic Interrelationships, 1910-1953. By FLANDERS DUNBAR. New York, Columbia University Press, 1954. Pp. xxii, 1192.

This huge volume is the fourth edition of a work which first appeared in 1935. It is an important source-book for all who are concerned with the relationships between psychological and physiological processes, whether from a theoretical, an experimental, or a clinical standpoint, although it is designed primarily to educate the physician to the importance of psychological processes. In 751 pages of text, the author abstracts, summarizes, quotes (many times at length), and cites 4,717 numbered references and about 500 additional unnumbered references. This monumental accomplishment is intended to provide only a key to the literature—a starting bibliography!

The book is far more than a survey of emotions and bodily changes in the traditional sense of academic psychology. It is an heroic interdisciplinary effort to comprehend the mechanisms of "intra-organismal equilibrium" and the "equilibrium of organism-environment," with their implications for health and disease, adjustment and maladjustment. Experimental psychology, psychiatry, biology, physiology,

philosophy, psychoanalysis, mathematics, statistics, engineering, zoology, medicine, sociology—all are represented, albeit with varying detail, expertness, and bias. A serious and scholarly attempt is made to interweave all disciplines relevant to the problems of psychosomatic medicine. The weave is necessarily loose, the colors clash, and the final garment has graceless lines; but it is a garment, and future tailors of science may make it more beautiful and more functional.

The book will not answer the challenge of those psychologists who think in purely behavioral terms and prefer to remain in ignorance of the events beneath the skin, nor will it even begin to satisfy the curiosity of those who wish to comprehend the organic mechanisms 'underlying' psychological function. It will, however, illuminate and clarify in exciting detail the bodily sequels of the storms and stresses with which the organism must cope.

The book is not a dispassionate one. It is frankly evangelistic, bringing the light of the Psyche to those engaged only in crass commerce with the Soma. 'Either-or' formulations are the work of the devil. Conversion involves only a willingness to accept and expand the notion that there is high scientific and therapeutic utility in simultaneous investigation of psychological and physiological processes. The high priest of the religion is Freud, but his words do not appear frequently because, in Dunbar's view, the contributions of psychoanalysis are scattered and incidental, and because only the psychoanalytically trained reader can appreciate them. No doubt is left, however, that the elaboration of psychoanalytic theory and practice is the royal road to psychosomatic medicine.

The reader who has a more temperate evaluation of the contribution of psychoanalysis, of the validity of the Rorschach test and finger painting, of the effectiveness of therapy, and of the experimental basis for many broad generalizations, will be exasperated often, but he will be fascinated by literature he has never read, implications he has not seen, relationships he has not detected. He will be exasperated by the repetitious testimonials of surgeon, internist, neurologist, and public-health official to the impact of psychosomatic medicine, but he will be rewarded by the new avenues of thinking uncovered. He will be irritated by some extremely glib statements (e.g. "the child who takes his first breath easily and normally is not so likely to be a feeding problem as is the child who struggled to get his first breath"); but he will be thankful for the subtlety of the treatment of 'mind-body relations' (e.g. in the discussion of neurological disorders: "our understanding is inevitably incomplete so long as we consider the mental disorder the necessary and self-evident result of a brain lesion, failing to consider the influence on the result of the organic lesion . . . exerted by the personality and previous experience"). He will detect non-sequiturs, awkward transitions, blind spots, and unnecessarily long polemics; but he will also note scores of papers he must look up because they are directly relevant to his work and his thinking.

The book is organized in three parts: (I) Orientation and Methodology, (II) Organs and Organ Systems, and (III) Therapeutic Considerations and Concluding Remarks. There are, unfortunately, three separate bibliographies. The index is 175 pages long, and a joy. It is compulsively detailed, and liberally cross-referenced. The most minor point can be tracked with great thoroughness.

Even to list the concepts developed by the author would be impossible here. The main guiding abstractions are *intra-organismal equilibrium* and *organismal-environmental equilibrium*. The emphasis is on dynamic rather than static measurement,

on longitudinal rather than cross-sectional approaches both to behavior and to disease, on temporal sequences of "psychosomatic functions" rather than on parallel or "expressive" sequences. The last paragraph of the penultimate chapter excellently summarizes the nature of the material to be found in this book and the justification for the appearance of a fourth edition. "The literature of the three to four decades preceding 1933 is helpful primarily in questions raised. The bulk of the material cited or abstracted in the sections added to each chapter in Parts One and Two provides more convincing evidence that soma can be influenced through psyche in ways increasingly subject to psychotherapeutic control. Study of the literature indicates also that appropriate application of such techniques as suggestion, hypnosis, briefer psychotherapy, and psychoanalysis is of fundamental therapeutic significance in prevention of disease. They are helpful in early diagnosis and treatment and in the cure of certain diseases where physiological changes are reversible. Psychotherapy is useful also as an adjunct when serious structural changes have occurred. Since it improves response to general medical procedures and may eliminate at least one stimulus to further unfavorable change, it is useful even in the treatment of illnesses where a fatal outcome is to be expected. Why this is true is becoming clearer, although many formulations about what psychophysiological mechanisms are involved remain untested."

Fels Research Institute   /   JOHN I. LACEY

Introduction to the Psychology of Music. By G. RÉVÉSZ. Translated by G. I. C. DE COURCY. Norman, University of Oklahoma Press, 1954. Pp. xiv, 261.

The present volume is an English translation of Révész's *Einführung in die Musikpsychologie* which appeared in 1946. In a preface to the English translation, Révész describes the psychology of music as a new field, although he makes use of sources which indicate that the field is about as old as psychology itself. Neither is the field new in the literature available in English. We have only to think of Ellis' translation of Helmholtz, and the publications of Max Meyer, C. E. Seashore, Max Schoen, C. C. Pratt, and we realize that there is a literature in English which offers variety in subject-matter and point of view. It is true, however, that de Courcy has rendered a valuable service in making Révész's book available in English translation, because the background, methods, and theoretical orientation of this work are somewhat different from those of the English works in this field.

The book, not a large one, is divided into three main parts. The first, entitled "Hearing, Sound, and Musical Tone," contains brief, incomplete discussions of the psychophysics of tone, tunings of the scale, auditory physiology, and the vocal production of tones. The second, on "Fundamental Problems of the Psychology of Sound," discusses the hearing of musical tones, intervals, consonance, absolute and relative pitch, and one or two other questions of musical perception. The third part, "Fundamental Problems of the Psychology of Music," discusses some features of musical talent and its pathologies, and concludes with a few remarks on the origin of music and a few more on aesthetics.

Révész believes that his own theoretical and experimental contribution to our knowledge of the perception of musical tones and intervals is especially important, and in the second part of the book he expounds his two-component theory. As far as musical perception is concerned, he says, the identity of a tone is composed of its

pitch and its tonal quality. In this case "tone quality" does not mean timbre, but the position of the tone in the octave-cycle, considered without regard to pitch. As far as this quality is concerned, all Cs are identical, though any two Cs might differ from one another in pitch by some number of octaves. According to the theory, a map of tonal experience can be drawn as a rising spiral, the upward dimension expressing pitch, and each completed cycle of the spiral representing all the tones within the compass of an octave. The spiral may be thought of as inscribed on the surface of a vertical cylinder, and any tone can be defined in terms of its position on the surface of the cylinder, at some specific height and some specific direction from the axis. A vertical line on the surface of the cylinder intersects the spiral tone-line once every octave, and these intersections will define a set of tones all identical in quality but differing in pitch by some number of octaves.

Révész calls this graphical method of representing musical pitch, a theory referring to a body of controversial literature in which he has defended it against attacks by Stumpf and others, and disputes the credit for authorship with Hornbostel. It is really only a descriptive device, and should be regarded as a piece of rhetoric. That the description is true can be settled by each person who refers to his own musical experience. Its truth, however, does not deny the truth of any other rhetorical device, and neither does it describe all of tonal experience. It is a poor description of musical experience compared to Max Meyer's descriptions of the experience of tonality, which Révész ignores. He might have used Meyer's description in order to have a phenomenal content for the angular dimension of his graph.

Révész values his theory because of its useful relation to other problems. He says there are cases of pathology of musical perception in which only one of the two components is erroneously perceived. He cites one case (is it the only case known?) in which a patient, could discriminate pitches correctly at one sector of the pitch range, but erroneously assigned the quality of G-sharp to the six tones from A to D in that particular octave, resuming correct identification again at D-sharp. This finding was reported by Révész and Liebermann (who was also the patient). A second reason given for the value of the two-component theory is that it is a necessary prerequisite for explaining the perception of intervals. The perceptual characteristics of an interval are derived from the pitch and quality of each of the notes that compose it. This review cannot deal in detail with Révész's description of musical intervals, but the prospective reader may be guided by one additional remark: Intervals that are physically the same on the tempered keyboard are treated as perceptually the same, e.g. an augmented fourth and a diminished fifth.

A spot-check of the translation against a page of the original shows that it is good. The sense of the material is accurately transposed into English that is flexible and readable, and, for the most part, free of contamination from the sentence-structure and other language-patterns of the German text.

Although this review may seem to be unfavorable in its analysis of the two-component theory, the book has a place in the literature in English on this subject, and it is one of a number that anyone interested in the psychology of music should read. It holds its own in comparison with other such books, some of which fail to consider music as phenomenal content, approaching the subject as if music were nothing more than the frequency, amplitude, and complexity of sound-waves.

University of Florida HENRY WUNDERLICH

Mathematics and Plausible Reasoning. By G. POLYA. Princeton, Princeton University Press, 1954. Vol. I, *Induction and Analogy in Mathematics;* Pp. xvi, 280. Vol. II, *Patterns of Plausible Inference;* Pp. x, 190.

George Polya is an eminent mathematician and teacher of mathematics. These volumes constitute, in many ways, a case-history of how he works in mathematics. He attempts to describe how he and other mathematicians actually think about a problem, how they guess at an answer, how they get hunches about true theorems, and incidentally how they prove them. These steps in thinking stand in sharp contrast to the logical structure of a proof which is imposed after the main thinking is over. One might almost consider Polya's book a 'confession.' His introspection is at all times candid and often humorous.

The essential steps used by mathematicians in thinking about their problems are present in the thinking of scientists in other fields. Psychologists may find Polya's work valuable, not only for its application to their own thinking, but as a source of suggestions for research on higher mental processes. The work also should be of interest to students of mathematics, for Polya tries to put into words what the student ordinarily learns only by serving an apprenticeship.

Volume I is called "Induction and Analogy in Mathematics." Without doubt, inductive reasoning is at the core of scientific thinking. One conjectures a general relationship by examining very limited data and then looks for ways to test that conjecture. Without induction, science would come to a halt, but Polya warns of the dangers involved. One of his many examples is the famous sequence of numbers $(2^{2^n} + 1)$ which Fermat conjectured led only to prime numbers because it gave primes for $n = 1,2,3,4$; Euler later found that the term for $n = 5$, divisible by 641, is not prime.

When a proposition which is true for a given set of objects is extended to a larger class, *generalization* occurs, says Polya. The reverse process he calls *specialization*. Formal characterization of these processes is straightforward. *Analogy,* however, is a more difficult concept, for it leads to problems in defining "similarity," problems that are well-known to psychologists. Nevertheless, analogy is a powerful tool of the scientist, not to prove theorems but rather to make discoveries which can later be proved or disproved by other methods. For example, Euler discovered by analogy that the sum of the series, n^{-2}, for positive intergers n, is equal to $\pi^2/6$.

Polya provides a wealth of examples of how induction and analogy have led to the solution of problems in infinite series, geometry, and number-theory. In all these illustrations, induction leads to conjectures, not proofs. In Chapter VII, however, the author describes the method of *mathematical induction* which is used to demonstrate that a conjecture is true. The balance of Volume I is devoted to more involved mathematical problems and the way in which induction and analogy have aided in their solutions.

Volume II, called "Patterns of Plausible Inference," is an attempt to formalize certain basic aspects of scientific thought. It is an extension of logical calculus to include "plausible calculus." For example, by the rules of ordinary logic, we conclude that A is false if (1) A implies B, and (2) B is false. On the other hand, if A implies B, and B is true, formal logic says that A may be true or may be false. Polya says that the conclusion actually made by scientists in the latter case is "A is more credible." Moreover, if B is very unlikely, but turns out to be true, then A is "much more credible" as a result. This is a pattern of plausible inference which has

no place in formal logic but is an integral part of scientific thinking.

The pattern just mentioned is that of examining a consequence. Another pattern of plausible inference is examining a possible ground: if A is implied by B, and if B is false, then A is "less credible." Still another pattern is that of examining a conflicting conjecture: if A and B are incompatible, and if B turns out to be false, then A is "more credible." These patterns, their combinations, and their extensions appear to cover most of the culturally accepted steps in scientific reasoning.

Polya is next led into a discussion of some fundamental ideas in probability-theory and statistical inference. In this area, Polya makes no new contribution (nor does he intend to), but he does discuss the subject in a new way. The reader who is not completely comfortable with modern statistical thought may find Polya's treatment rewarding, provided he has the patience to follow the author this far.

The book is concluded with a chapter on plausible reasoning in invention and instruction. Throughout the two volumes there are a great many examples and problems, the solutions for which are given at the end of each volume.

A few words should be said about the level of mathematical training necessary for reading Polya's two volumes. For complete understanding a year of calculus is adequate, but even this is not essential for following the main arguments. The examples involve more mathematics than does the main part of the text. The psychologist who lacks confidence in his own mathematical skills should have no serious handicap in reading most of the first four chapters of Volume I nor in reading nearly all of Volume II.

Investigators of higher mental processes will be most interested in the first few chapters of Volume II. To this reader, at least, they suggested a number of possible experiments to measure the increase in "credibility" of an assertion as new evidence is given, and to determine whether or not Ss do in fact employ Polya's patterns of inference.

Harvard University   ROBERT R. BUSH

The Origins of European Thought about the Body, the Mind, the Soul, the World Time, and Fate. by RICHARD BROXTON ONIANS. Cambridge, The University Press, 1951. Pp. xvii, 547.

The book begins with an etymological discussion of several terms in the *Iliad* and the *Odyssey* which refer to bodily organs and their functions. On the basis of traditions and customs collected from regions widely separated, the author proposes a number of startling assumptions and ends by discovering a highly complicated system of psychological ordering. Into this texture fit many details known about Egypt, Crete, Greece and Rome, the Near East, and Nordic as well as Celtic culture. Consistency is likewise obtained for customs traced back to the neolithic and even the paleolithic period. The evaluation of the etymological material must be left to others. The question of how far consistency supports a historical reconstruction also lies beyond the scope of this review. For the psychologist, Onians' book is important because it opens a view into the prehistoric beginnings of psychology.

This is not a psychological text. Although the author endeavors to relate the ancient thinking to modern psychology, the reviewer finds it more profitable to report the psychological information of the text in a terminology and a frame of reference to which she is used. The terms that express reactions are less differentiated than those of today. If we equate them with the verbs 'to perceive,' 'to know,'

'to think,' and so forth, we do not render justice to their complexity. The conceptual analysis into perceptual, cognitive, emotional, conative phases, which somehow still hinders a descriptive approach, was foreign to the early period. It seems, therefore, that the term 'total reaction' would be closer to the early use of verbs which later carried specific meaning.

The same complexity extends to the stimulus-objects, which are characterized as angry, dear, honest, and the like. This usage would imply a reaction to the so-called physical as well as to the physiognomic qualities of the stimulus. Stimuli are considered to be of one kind, and although they enter through different receptors they all reach the same organ. In contrast to modern psychology, which until recent times has neglected speech, this ancient psychology starts from speaking. Words are identified with breath, and breath is concrete. The breath comes from one person and goes to another. It enters through the ears and reaches the breast, in particular the lungs in the breast. Those who have received breath are wise. Breath may likewise enter through the eyes and the other receptors. It is the physical stimulus for all sensory organs, and it always goes to the lungs. Prehistoric and early historic times lacked the anatomical and physiological information which speaks against such a generalization. On the other hand, it seems rather modern that all perceiving is considered as active rather than passive. The eyes, for example, do not actually 'receive' the breath, but they 'breathe it in.'

There are two kinds of behavior and two corresponding controlling agents. This is the old distinction between voluntary and involuntary behavior, more recently operant and reflex behavior. The two agents are called 'thymos' and 'psyche.' The first—unlike the gland with which it has nothing but the name in common—is located inside the lungs and the heart. The 'thymos' is something concrete. It is vaporous, blending and interacting with the air outside and at the same time related to blood. The blood makes it moist and warm. At the death the 'thymos' leaves the body and suffers destruction. All vital processes and all behavior which is not reflex are controlled by the 'thymos.' It depends on food and drink. In later times the 'thymos' lost its concrete character, developed into something abstract, namely emotion, and more specifically, rage.

The 'psyche' is located in the head. Actually, however, neither agent is strictly confined to the one region which is its seat. The 'psyche' can influence the 'thymos' for good or for bad. It knows the future. Reflexes like sneezing tell us whether something said is true or whether it is going to happen. A shudder down the spine is a premonition of death. The 'psyche' has nothing to do with ordinary awareness. It is a vehicle of life which continues after death. At the same time it is a starter of new life and it is also that part of us which is holy. The actual length of lifetime is determined by the 'aiōn,' the liquid in the body, especially identified with the cerebrospinal fluid. It contains the gaseous 'psyche.' In the course of history the 'psyche' became enriched by taking over some of the functions of the 'thymos.' Thus started the dispute whether the governing power is in the head or in the breast. Although psychologists are no longer interested in the differentiation of mind and soul, they still speak of control in linking up specific brain regions with particular behavior.

The short review does not do justice to the complexity of the concepts and the intricacy of the psychological system which brings good order into a surprisingly great number of observations on human behavior.

Pacific University

ANNA BERLINER

Madison Butler

Madison Bentley: 1870-1955

Coöperating Editor 1903-1926; Co-Editor 1926-1950

Madison Bentley, psychologist, author, editor, critic, and teacher, one of the first generation of psychologists trained in America, died in his eighty-fifth year in Palo Alto, California, during the early morning of May 29, 1955. His death, which followed an illness of several months, was due to a series of coronary occlusions. He was born in Clinton, Iowa, on June 18, 1870, the fourth child and second son in a family of eleven children—five boys and six girls.

He was christened Isaac Madison, family names from both sides of his house. He abbreviated his first name, however, sometime during his early manhood[1] and thereafter until 1909, when he dropped the 'I.' "because it was often misprinted 'J.',"[2] which was always the case, of course, in German periodicals and references.

Genealogy. His father, Charles Eugene Bentley, a descendant of English stock that settled in Massachusetts during Colonial days, was a Baptist minister, born and reared on a farm near Warners, a village in central New York, northeast of Syracuse. Charles Bentley was educated in the available country schools and the academy in Cazenovia, New York. It was at the Academy that he acquired his pervading interest in history, literature, and politics which led him to head the Liberty Party and to accept its nomination for the Presidency of the United States in 1896.

After completing his formal education, Charles married Persis Orilla Freeman, a neighbor's daughter, in 1862. She was also well educated, particularly for a young woman of her time. After common school, she completed her training at Falley Female Seminary and then taught school until her marriage. She had a flair for writing and directing plays and entertainments and was, as one of her daughters described her, a "capable, romantic, sensitive, possessive, hard-working woman."

---

[1] His first publication in 1897 (see accompanying bibliography) was under the name "I. Madison Bentley."

[2] Probably a rationalization concealing the real reason; see Footnote 3.

Hearing and heeding the call to the ministry, Charles, soon after the birth of his first child, a daughter, moved his little family to his first pastorate in Clinton, Iowa. He remained there for thirteen years, during which time his next six children, a girl, two boys, and then three girls, were born. In 1878 he moved his growing family, "to get away from the city atmosphere and the influences of a river town," to a pioneer's life on a farm in Butler County in eastern Nebraska, twenty miles from the county seat and nine miles from the nearest post office. Here he farmed and took charge of a large and scattered flock. He married, baptized and buried his parishioners, advised them in political and financial matters, doctored them with the aid of a medical book during their illnesses, and was in general the intellectual as well as spiritual leader in his part of Nebraska.

Education. It was in this setting and atmosphere that the young Madison grew up. His boyhood was spent working on the farm and in acquiring the solid rudiments of an education which he obtained at home and at a one-room country school that "never lacked a good teacher from the east."

In 1884, when fourteen years old and ready for his secondary education, which he could not get in a country school, Madison and his older brother, Rufus Clarence, left the prairie home for three profitable and happy years with a wealthy uncle and aunt in Brooklyn, New York.[3] Being childless, these relatives treated the boys as their own and gave them all the advantages of life in a big city. Both boys attended Adelphi Academy (now College), where they enrolled in the college preparatory course with the idea of continuing their education in an eastern university. Madison in addition studied music (organ and piano) under a gifted teacher, who inspired him to apply himself so assiduously that he became quite proficient on both instruments. A musical career was a possibility, but this and his dreams of an eastern education were abruptly shattered during the spring of 1887 when he and Rufus were, over their protestations, summoned home by their father that other members of the family could enjoy the advantages of their uncle's generosity and hospitality.

During Madison's absence in Brooklyn, a new railroad had been constructed, and the village of Surprise had sprung up upon it, two miles from the Bentley home. Just as he returned, a bank was being built in this little village and, though only seventeen years old, he was offered the position of cashier which he accepted. It was an escape from the farm—and, to prepare himself for the work, he spent the ensuing summer at the Omaha Business College studying banking. In the following fall, he assumed the full responsibilities of the position and carried them for the next three years.

In 1889, Charles Bentley left the farm for a pastorate in Lincoln, Nebraska, that his rapidly maturing family might enjoy the educational advantages of the city and the State University located there. Madison, however, preferred to remain at the bank. He had, in emulation of his uncle, decided upon a financial career;

[3] This uncle, James Madison Ham, was the first treasurer of the Union Pacific Railroad and subsequently a member of the firm of John Dillon, later Dillon, Reed and Company. It was from this uncle that Madison received his second given name, hence it is highly probable that 'Isaac' was first abbreviated and then dropped either as an indication of his gratitude to his uncle or as an attempt to identify himself more closely with this highly successful relative, something a reticent and shy young man, such as Madison was, could neither express nor openly admit.

gone were his musical and scholarly ambitions; he had now neither time nor desire for a university education.

His father's wishes, however, again prevailed. He rejoined the family in Lincoln in the fall of 1890, and sought to register at the University. Despite his three years at Adelphi Academy, he found that he did not have sufficient credits for entrance, but he was permitted to register as a special student to make up his deficiencies.

He qualified in one year and entered the University as a regular student in the fall of 1891. By then he had discovered that university life was exceedingly stimulating and he entered enthusiastically into all phases of it. He joined a college fraternity (Phi Kappa Psi) and threw himself into the activities of the University, but soon he found that his greatest interests were in the classroom. His first enthusiasm was for chemistry—upon which he concentrated during his probationary year; then his interest shifted to the military—in which he served as adjutant of the University Cadet Corps under the Commandant, Capt. John J. Pershing, later commander of the U. S. Expeditionary Forces in World War I, with whom he formed a lasting friendship. Then he turned to English literature—being entranced by Browning, many of whose poems he committed to memory and was able to quote even in the declining months of his life. Geology and its allied fields of paleontology and anthropology then claimed his attention. These interests carried him on two exploratory expeditions digging in the sandhills of western Nebraska.

These subjects, the effects of which may be seen throughout his life in his thoughts, activities, and style of writing, yielded finally to an enthusiasm for psychology during his first course with Wolfe,[4] whom he described in his autobiography as "a very engaging man"[5] and later acknowledged as his professional mentor.[6]

An assistantship in psychology under Wolfe during his senior year, his election to Phi Beta Kappa and an offer of a scholarship in psychology from Cornell University weaned him from his other academic interests and also away from a financial career, a smoldering desire for which he had maintained throughout his college course by working during vacations and spare time in the banks of Lincoln.[7]

---

[4] H. K. Wolfe and J. McKeen Cattell took their doctorates at Leipzig in 1887, the first two of a long line of American psychologists to be trained by Wundt. See M. A. Tinker, Wundt's doctorates and their theses, this JOURNAL, 44, 1932, 630-632.

[5] Madison Bentley, *A History of Psychology in Autobiography,* Clark University Press, 3, 1936, 58.

[6] Though Bentley received his Ph.D. degree at Cornell University, he acknowledged Wolfe, not Titchener, as his chief professional mentor. See M. D. Boring and E. G. Boring, Masters and pupils among American psychologists, this JOURNAL, 61, 1948, 528.

[7] The biographical data given here were obtained from his youngest sister, Laura Persis Bentley Hall; from his second wife, Margaret Russell Bentley, who was able, with the aid of an old diary kept by Uncle Ham (see Footnote 3) to confirm the facts reported; and from his daughter, Virginia Bentley Meehan, who supplied the information about her mother. Madison was never one to talk about himself, his family, or his early life and experiences. Regarding such matters, he was strangely reserved and reticent. Though closely associated with him for forty-odd years, the writer knew practically nothing about any of the items just reported. Except for date of birth and in a few instances the names of his parents, he omitted all other information about family from his listings in *Who's Who* and other similar compendia.

Whatever the determinants of his choice of a career in psychology were, the fall of 1895 found him at Cornell University. The early years there "proved less illuminating and absorbing than the collegiate years had been."[8] Titchener, a a young Englishman direct from Wundt's laboratory and a recent addition to the Cornell staff, was then an unknown who had yet to win his spurs. He and Bentley were, moreover, about the same age. Under these circumstances it is hardly to be expected that Titchener would replace Wolfe, an older man who was also a Wundtian product, as Bentley's 'father figure' in psychology. Indeed, Bentley regarded Titchener, then and throughout his life, more as a fellow student and colleague than as a revered professor or an intellectual mentor. The Cornell system of graduate training, a transplantation from Germany, was also responsible in large measure for this attitude. No graduate courses were offered and none of the undergraduate courses was required. The graduate student was allowed freedom to prepare himself for the final examinations in the major and minor subjects of his choice in any way that he wished: by studying in the library, experimenting in the laboratories, or attending undergraduate courses.[9]

As a result of this freedom, Bentley drifted during his first year at Cornell and acquired, as was his wont, a new enthusiasm—now it was for philosophy. In this he was encouraged, as he writes, "by great liking and respect for my preceptors in those studies [Cornell had an outstanding faculty in philosophy], and also by a lack of satisfaction in the local views and practices within my primary field of research."[10]

This philosophical excursion, which yielded his first scholarly publication,[11] was brought to a close by the possibility of an assistantship in psychology. Pillsbury, who held the position, received a call from Michigan and was leaving at the end of the academic year. To advance his own candidacy for the position, which if obtained would make his marriage possible, Bentley applied himself to his work in psychology and soon found that the enthusiasm with which Titchener taught and carried on research was "so infectious" and "the air of the laboratory was so charged with topics and problems of research,"[12] that the logical side of the training, despite its emphasis on concept and system, to which he had previously objected, was of immense value to him and his fellow students, among whom were Washburn, Pillsbury, Gamble, Bagley, Whipple, and Baird—an outstanding group, every one of whom was destined to make a mark in psychology.

Bentley applied himself so zealously to his dissertation and participated so ardently in the other activities of the department that he was appointed to the coveted assistantship in June 1897. The following August he married Emma Snelling at Marshalltown, Iowa. She was a graduate of Grinnell College and a talented musician, to whom he had became engaged during his undergraduate

His autobiography, in which one normally expects to find such information, was of no help, as he did not give even the date or place of his birth in it.
[8] He discusses the reasons for his choice of psychology at length in his autobiography, *op. cit.,* 55-57.
[9] Many students are unable to survive this degree of freedom, but those who do, those who learn to depend upon their own initiative, are likely to continue as productive scholars long after they obtain their doctorates.
[10] *Autobiography,* 57 f.
[11] The psychology of 'the grammar of science,' *Philos. Rev.,* 6, 1897, 521-528.
[12] *Autobiography,* 59.

days while she was an instructor in the Conservatory of Music at the University of Nebraska. One child, Virginia, was born of this union. During the following year he completed his dissertation, wrote a short article on psychology and education, and received promotion to an instructorship. He had embarked upon his life's work.

Career. In 1899 he published his dissertation on the memory image, a live topic in psychology at that time. This study had a large part in shaping the destiny of this concept.[13] He followed this research in 1900, with a study on *The Synthetic Experiment,* in which he analyzed and synthesized the perception of wetness. He found that it was an integration of pressure and cold. This study, a little classic, served as the model for many, similar studies in various laboratories. The following year (1901) he published a series of seventeen articles on *Illusions,* which appeared weekly in the scientific section of *The Chicago Record-Herald.* These articles, which covered the full range of the subject, are excellent; they deserved a more permanent and accessible medium of publication.

In 1902 he was promoted to an assistant professorship, the rank he held until he left Cornell in 1912. Not only were promotions in rank made slowly in those days, but in addition Cornell had, at that time, only three academic ranks: instructor, assistant professor, and professor;[14] hence promotions to a professorship were usually long delayed.

In 1903 Bentley joined the coöperating board of editors of this JOURNAL at the invitation of G. Stanley Hall; this was the first of numerous editorial positions which he was to hold, and the one that he maintained the longest. Near the end of this decade he began work upon his *Manual of Comparative Psychology,* which was available in mimeographed form for the use of his students in 1911.

In 1910 when Titchener was elevated to a graduate professorship and relieved from all undergraduate instruction—a promotion that was made to meet the offer of a graduate professorship from Clark University—the Department was divided and Bentley was made chairman of the undergraduate division and director of the undergraduate laboratories. Bentley might have advised with Titchener about the management of the undergraduate department but, since the separation was made in accordance with Titchener's wishes, he did not do so. He was charged with the responsibilities of the department and he exercised them. Titchener learned of

---

[13] The memory image and its qualitative fidelity, this JOURNAL, 11, 1906, 1-48.
[14] The intervening rank, associate professor, had been abandoned by Cornell in 1897 and it was not reëstablished until 1937. Its reëstablishment was opposed by many members of the faculty because of the fear that it would, since it carried tenure, become a cul-de-sac.

Bentley's decisions on expenditures and appointments of instructors and assistants indirectly and only after they had been made.

From full control to no control of the undergraduate department, which he had built from scratch, was a contingency that Titchener had not anticipated. It is small wonder, therefore, that tensions developed and relations between the two men became strained. To clear the situation and to regain his control of the undergraduate department, Titchener vigorously supported Bentley's candidacy for the vacancy created at Illinois by S. S. Colvin's departure. Though loath to leave Cornell, Bentley accepted the call received during the summer of 1912 and joined the staff at Illinois the following fall as Professor and Head of the Department of Psychology and Director of the Psychological Laboratories.

He remained at Illinois for sixteen years. Soon after moving there, in the summer of 1915, his wife died. She had shared in his early trials and successes, had helped him with his apparatus, had served as an observer in his experiments (being one of the two principal observers in the synthetic experiment already mentioned) and had been of great help to him in his writing and in the preparation of his manuscripts.

It was during this period that Bentley became interested in phrenology and obtained half of the library and museum of phrenological casts and models from the Fowler Institute in Philadelphia for the psychological museum of the University of Illinois. His study in this field resulted in an article on *The Psychological Antecedents of Phrenology*. Soon after the publication of this paper, the United States entered World War I. Bentley immediately volunteered for service. He joined Knight Dunlap's organization in the Medical Research Laboratory of the Air Service at Mineola Field, New York, in November 1917 as a Captain. He was discharged as a Major in December 1918, but he continued his membership and interest in the Reserve Corps for many years. The work of the organization with which he served was centered upon the non-acoustical organs of the inner ear. The principal outcome of these investigations showed that the Bárány test used in the selection of candidates for pilot-training was worthless. He carried back to Illinois his interest in this field and for several years after his return significant studies were published from his laboratory by Coleman R. Griffith.

During the summers of 1921, 1922, and 1923 he taught at the University of California in Berkeley; and he spent the academic year 1922-1923 there on sabbatical leave while writing a textbook, *The Field of Psychology*, which was published during the spring of 1924. During this year of leave

he won himself a second wife, Miss Margaret Russell, a graduate student in psychology at the University of California, to whom he was married in June of 1923. No children were born of this union.

From the date of his second marriage, events of consequence in his life occurred in rapid succession. Gone were the days of quiet contemplation and study; he was cast into a stream of activities which required an eighty-hour work-week to keep abreast. In 1925 he served as president of the American Psychological Association. He also gave three lectures in the Powell series at Clark University[15] (the first two being critical expositions of structural psychology and the last an exposition of his own system which escaped the dilemma of the mind-body problem by regarding the organism as a single unit). Among other activities he started writing his book on *Man;* edited Vol. 31 of the *Psychological Index,* and accepted the editorship of the *Journal of Experimental Psychology.* He gave his presidential address on *The Major Categories of Psychology*—a critical review of the various schools of psychology 'promoted' at that time—at the annual meeting of the Association which was held at Cornell University during the Christmas holidays. While at this meeting he participated in the reorganization of the editorial board of this JOURNAL, whose editorship Tichener had just laid down. In coöperation with Washburn, Boring, and Dallenbach, he then agreed to take on the additional responsibilities of its editing.

In 1926 he began the arduous duties of editing two major psychological journals, one of which he continued for four years and the other for twenty-five years. Editing in those days was a very different task from what it has now become. Besides reading and evaluating manuscript, returning those rejected and editing those accepted, and handling personally all the correspondence involved, editors read galley and page proofs and saw the issues through the press. In addition to performing these duties, Bentley served as vice-president and chairman of Section I (Psychology) of the American Association for the Advancement of Science and brought his book, *Man,* to the stage where it was ready for its first mimeographed edition.

In 1927 he wrote two important, critical papers: one for the Washburn *Festschrift* on *Environment and Context,* and the others for the *Wittenberg Symposium on Feelings and Emotions* which bore the title *Is 'Emotion' More than a Chapter Heading?* In the same year he rewrote his book on *Man,* the first edition of which was far from satisfactory to him. In 1928, after teaching another summer at the University of California in Berkeley,

---

[15] *Psychologies of 1925,* Clark University Press, 1926, 383-412.

he went to Cornell University as Sage Professor of Psychology and chairman of the Department in succession to Titchener, who had died in August of the preceding year.

Regarding the impact of his service at Illinois, *The Alumni News* of that University, when reporting his death last year, published the following statement: "Leading members of the faculty point out that Professor Madison Bentley . . . did much to bring Illinois to its present position of greatness . . . he ranked with [the great teachers at Illinois, who were named] in bringing real distinction to the Illinois campus." This encomium after an absence of twenty-seven years—an interval during which most teachers are forgotten!

Shortly after he returned to Cornell, the psychological laboratories were extensively remodeled and the curriculum of the Department was completely revised—the first changes to be made in either physical plant or courses of study since his departure sixteen years before. Individual rooms for graduate students and specialized rooms for research (soundproof, constant-temperature, and visual) were constructed, a skilled mechanician was placed in charge of the machine shops, and the undergraduate laboratory was divided into many small rooms in which pairs of students could work undisturbed.[16]

During the spring of the following year (1929), Bentley was one of a small committee, meeting at Princeton University, which reorganized the Society of Experimental Psychologists. He was the first host and chairman of the newly organized Society at its next annual meeting held at Cornell University in the spring of 1930.

In *Psychologies of 1930,* published during the spring of that year, he presented, in an article entitled *A Psychology for Psychologists,* his particular point of view of psychology[17] which had been "briefly and imperfectly sketched in the *Psychologies of 1925.*"[18] In the fall of the same year he went to Washington, D.C., on leave of absence from Cornell, as chairman of the Division of Anthropology and Psychology of the National Research Council. His year at this post was a busy one, filled with many meetings and conferences. Of the various projects initiated during his term of office, one, the chairmanship of a special NRC Committee on Psychiatric Investigations, engaged him for several years thereafter.

The plan pursued by this Committee, which was supported by a grant

---

[16] For a description of the remodeled laboratory see this JOURNAL, 43, 1931, 295-300.

[17] *Op. cit.,* Clark University Press, 1930, 95-114.

[18] The psychological organism, *op. cit.,* 1926, 314-321.

from The Carnegie Foundation, began with a large number of conversations with various individuals actively and professionally engaged "in the psychiatrical arts of cure, in problems of commitment and classification, in the general problem of incidence, in the tasks of hospitalization, in the investigation of origin, cause, and nature of the disorders, in control and prevention, and in teaching and the executive control of psychiatric instruction."[19] Out of this primary reconnaissance came the publication of several articles, a book, and a trip abroad by the chairman to discover how the problem of mental disorder was attacked in England. The first article, an address by Bentley at the Mayo Foundation in Rochester, Minnesota, on *Mind, Body, and Soul in Medical Psychology*, was published in 1933. During the following year (1934) the book, *The Problem of Mental Disorder*, appeared. It was written under the auspices of the Committee by twenty-five different authors and edited by Bentley and Cowdry. It is divided into four sections. In the first, Bentley outlined the problem; in the second, five leading psychiatrists expounded their own points of view or systems of treatment; in the third, twenty experts, chosen from the field of the natural sciences for their ability in productive research, suggested ways and means of advancing the psychiatrical arts; and in the fourth, Bentley gave his comments and reflections on the various contributions. The book is one that should be read by every person interested in the problem of mental disorder. Bentley's trip to England during the summer of 1937, which was made possible by an additional grant from The Carnegie Foundation, resulted in an article on *The Problems of Mental Disability in England*.[20]

During this excursion into the field of mental disorders, Bentley was involved in many other activities. Besides his administrative, teaching, and editorial duties, he wrote his autobiography, revised his textbook, worked in New Mexico with the American Indians during the summer of 1935 as research specialist of the U. S. Department of the Interior, taught during the summer of 1936 on the West Coast, directed numerous graduate dissertations, collected and edited the *Cornell Studies in Dynasomatic Psychology*, began his long series on The Research Committee of the American Otological Society, and published numerous articles, foremost among them being his contribution to the Golden Jubilee Volume of this JOURNAL on *The Nature and Uses of Experiment in Psychology*.

The revision of his textbook, which was completely rewritten, appeared

[19] Madison Bentley and E. V. Cowdry (eds.), *The Problem of Mental Disorder*, 1934, v-vi.
[20] *Op. cit.*, this JOURNAL, 1938, 1-17.

in 1934 under the title *The New Field of Psychology*. It presented the dynasomatic system of psychology anticipated in his article in the *Psychologies of 1930*. His autobiography—published in *A History of Psychology in Autobiography* in 1936—is an odd contribution. As an historical résumé of the derivation of his system of psychology, it is excellent, but as a biography it is sadly wanting. No biographical data—nothing about his parents, himself, or his family—are given within it, not even the place or date of his birth.

In 1935, the fortieth anniversary of his baccalaureate, the honorary degree of LL.D. was conferred on him by his Alma Mater, the University of Nebraska. In June 1938, when he had reached his sixty-eighth year, he was retired from Cornell as Emeritus Professor of Psychology. He was at the height of his productivity at this time, still possessed with excellent health and physical vigor; hence he was far from ready to lay down his academic duties.[21] The rules regarding retirement were, however, inexorable—no exceptions could be made.

To mark his retirement, his students and colleagues of the universities of Illinois and Cornell, at a dinner given in his honor, presented him with an oil portrait of himself painted by Olaf M. Brauner (a reproduction of which appears as the frontispiece of the July 1938 number of this JOURNAL) and a book of letters.[22]

Bentley's services to psychology did not end, however, with his retirement. He immediately joined the staff of the Library of Congress in Washington, D.C., as consulting psychologist—a position he held until 1940 when Herbert Putnam was retired as Librarian and all consulting positions on the staff were abolished. He then went to Palo Alto, California, to make his future home near his daughter's, who lived in the Bay Region. While there he was given the courtesy of an office in Stanford University and continued his editing and writing under the auspices of that University.

He did not, however, stay there for long; he was far from ready for the shelf of retirement. During the spring of 1941 he taught at Goucher College in Baltimore, Maryland, and in the fall of 1942 he returned to Cornell University as Lecturer in Psychology to replace Dallenbach, a

---

[21] During his last year at Cornell he wrote six articles and directed the doctoral dissertations of five students.

[22] For some peculiar reason he loathed having his picture taken and few indeed are the photographs, or even 'snap-shots' of him. It was, therefore, an accomplishment of the first order to get him to sit for the painting of his portrait. The combined efforts of his wife, the artist, and his closest friends were required to overcome his prejudices against it.

reserve officer from World War I, who had been recalled to active duty. In addition to teaching his old courses at Cornell, he took over the managing editorship of this JOURNAL and shortly afterward, when the editor of the Book Review Department entered the Service, he also took over the responsibilities of that Department. Though copy was short and reviewers were scarce, because of the large number of psychologists in the Services, he managed to issue the numbers of the JOURNAL in full size and on the dates due.

He carried the managing editorship for three years (1943-1945), until Dallenbach returned from the Service, and the book-review editorship until he retired from the JOURNAL in 1950 to devote the remaining years of his life to his own writing.[23] During his managing editorship, he began a series of timely and highly critical articles on *Tools and Terms in Recent Researches* and added a section to the Book Review Department on *Brief Comments on Recent Books*. His "Tools and Terms" inveighed against the shoddiness of many of the concepts and expressions in current psychological writings which resulted from contamination with the arts of practice during the War; and his "Brief Comments" were terse statements, all from his pen, about books cognate to psychology.

After his war-time duty at Cornell, he returned to the West Coast to make his home. He went first to Pasadena where he was given office space in the California Institute of Technology to continue his editing and writing but after one year there he returned to Palo Alto to be nearer his daughter. Here again Stanford University accorded him the courtesy of an office, which he used regularly until his last illness.

Achievements. Bentley's genius lay in his erudition and critical acumen. He had no distracting hobbies nor interests outside of psychology, which was both his vocation and avocation.[24] He was, consequently, widely read and conversant with the current as well as with the older literature in psychology. His professional life was practically coextensive with psychology in America. He saw movements wax, flourish, and wane. He was the observing critic and he criticized freely in his reviews and addresses. He regarded the reviewing of books—pointing out shoddiness as well as excellence—as among the important duties of a scientist, and he gave

---

[23] An editorial announcement, this JOURNAL, 64, 1 f.
[24] Bentley never lost his interest in nor appreciation of music though he never, during the writer's acquaintance with him, deigned to perform upon either piano or organ. During his younger days at Cornell, he played a good game of tennis—due principally to his ability to play with either hand—to shift his racquet from one hand to the other as occasion demanded. He was ambidextrous; wrote and used tools and other instruments equally well with either hand.

unsparingly of his time and energy to this task. Over 250 reviews were written by him; many were published under pseudonyms and the names of his students that his prolificacy would not be apparent—but his 'signature,' in his style and use of words, was always upon them. As a critic in psychology he was without an equal.[25] We shall not soon see his like again.

(a) Books. Bentley's genius was also his greatest handicap and weakness. He was so self-critical and so keenly aware of the gap between his own writings and the ideal which prompted them that he brought but two of the various books written during his long and productive life to publication. His first book on *Illusions,* mentioned above, was published serially in 1901 in the scientific section of a Chicago newspaper.[26] It was, to be sure, published at a time when numerous authors, especially in Europe, were writing about illusions,[27] and it is possible that the series contained nothing new psychologically, but it did indeed bring together all the illusions and facts then known and it was without doubt the best series of articles upon the subject available in English. Since it was published in a newspaper, it was not indexed in any of the scientific bibliographies, hence the series had no impact upon psychology and was completely lost. Publication in book form would have saved it—and a book at this point in his career, when he was still a young instructor, would have added to his stature in psychology. Despite all practical considerations, he did not collect and publish the series because it fell short of the perfection that he demanded of himself. Thus his first book went by the board.

His second book, *A Manual of Comparative Psychology,* which was used in mimeographed form by the writer in 1911, was written at Titchener's suggestion. Shortly after the publication of his *Manuals of Experimental Psychology,* Titchener suggested to Bentley and Whipple that they, respectively, do for the fields of comparative psychology and testing what he had done for experimental psychology; namely, prepare manuals for them. The suggestions were timely, and both men immediately set to work upon their separate tasks. They completed their manuscripts by 1909. Their books were not perfect, but as first essays they

---

[25] Titchener once remarked to the writer that he regarded Bentley as the most able critic in psychology. In 1909, in inscribing his *Experimental Psychology of the Thought Processes* to Bentley, he wrote in the preface: "In dedicating the volume to Professor Bentley, I wish to express my gratitude for the help that he has generously rendered, not only in this particular case, but in all my literary undertakings of the past dozen years."

[26] For references see the accompanying bibliography.

[27] For references see *General Index, Vols. 1-30,* this JOURNAL, p. 121.

sufficed. Whipple brought his *Manual of Mental and Physical Tests* to publication in the spring of 1910, believing that imperfections could be corrected in later editions—and many were. Bentley, on the other hand, delayed publication; he sought perfection, as was his wont. He had his manuscript mimeographed for use in his classes with the idea of revising it as use and experience dictated. He thought that he could accomplish this end in one year, but just at that time the chairmanship of the undergraduate department at Cornell descended upon him and publication was again delayed. Before he could complete his revision he moved to Illinois, and a further delay was necessary. During this period, Watson launched his *Behaviorism,* which had to be taken into account; and the results of Pavlov's experiments reached America and could not be ignored. Revision followed revision, but events were moving so rapidly in the field of comparative psychology that it was impossible for him to keep abreast of them. The time for his *Manual* had passed; after a decade spent in its writing and rewriting, he finally laid it aside. Had he published it in 1910 and revised it in successive editions as occasion demanded, the course of psychology in America might have been very different. The rampant invasion of psychology by radical behaviorism, during the teens and early twenties, might have been tempered and perhaps avoided. Bentley's *Manual,* patterned after Titchener's, was worthy of its model. It is a great loss to psychology that it was not published in 1910.

Laying the *Manual* aside, Bentley turned to the field of the abnormal. Regarding his attempt at a book in this field, he wrote in his *Autobiography,* "Ill-considered plans to write a 'psychology of the abnormal' were checked just in time by the caustic remark of Adolf Meyer, whose counsel had been wisely sought, that he wished that *he* 'knew enough to write such a pretentious book'."[28] This book was also put aside but Bentley did not give up the idea. Years later, in 1934, he co-edited a book on *The Problem of Mental Disorder,* as already reported. If Meyer did not individually know enough to venture upon such a book, perhaps it could be accomplished by the conjoined efforts of the experts in the field. In any event, Bentley permitted himself to sponsor such a publication.

Bentley's fourth book, an elementary text entitled *The Field of Psychology,* published during the spring of 1924, was completed during the academic year 1922-1923 while he was on sabbatical leave. He was less self-critical in this work than with any other of his writings. The pressure of having a book to show for his year's leave, or the success of his second

---

[28] *Op. cit.,* 61.

courtship, which progressed simultaneously with the book's writing, or both combined, may have been responsible for the unusual lapse of his self-criticism; at any rate publication was not delayed.

Bentley sought, in this book, "to reconsolidate the disjoined parts [of psychology] and to redintegrate the old and the new,"[29] and also to treat of "the total living creature which is at once bodily (or somatic) and mental."[30] In Parts I and II of the book he describes the "constituents" of experience and their organization drawing heavily upon the content or structural schools of psychology. In Parts III and IV, he discusses the psychosomatic functions and the socialization and development of the psychological organism from a functional point of view. The two halves of his book stand logically and systematically apart; there is a chasm between them that can not be bridged and that he, at the time of his writing, either did not see or could find no way of avoiding.[31]

With the achievement of this book behind him, he turned to the writing of *Man,* which he intended to make a semi-popular book that considered *homo sapiens* in relation "to the earth and to the stars, and to the electrons and to the molecules." In this book his early anthropological training found expression. Having no urge for immediate publication, he brought all his erudition and critical acumen to bear upon it with the result that he revised it many times. He put it through four mimeographed editions for use by his students but, despite the urging of his publisher and closest personal friends, he would not permit its publication. It had not attained, in any of the four editions, the perfection that he demanded.[32]

His labors with it were not, however, entirely lost because he discovered during them that the logical gap in his *Field* was of his own making, that it could be avoided simply by "regarding the integrated body as the complete existential structure of the organism and by confining its primary descriptions to the psychological functions of the body."[33] These functions, five in number, are: apprehending (perceiving, remembering, imagining), executive (action, emotion), inspection, comprehension, and elaborative thinking. Their description and the description of their "conditions, government, and products, supplies a natural and adequate basis for psycholo-

---

[29] *Op. cit.,* vii.

[30] *Ibid.,* viii.

[31] That he was aware of the hiatus is shown by the fact that he wrote in the preface of *The New Field* that the earlier book was divided into "two irreconcilable parts" (*The New Field of Psychology,* 1934, vii).

[32] The writer used a copy of the third mimeographed edition (1929) as a reference in his classes until it was nearly worn out by handling. The students found it interesting and stimulating. Any one of the last three editions was worthy of publication.

[33] *The New Field of Psychology,* 1934, **vii.**

gical growth and development, as well as for the social aspects of man's activities, and for all the defects and disorders of a psychological kind."[34]

In common with the founders of psychology, Bentley regarded the subject as a separate, independent science. He believed, moreover, that its progress, which had been promising up to the first decade of this century, had been interrupted, delayed, and corrupted since that time by the repeated invasions of other disciplines and the arts of practice which referred their psychological problems to it for a quick solution. Psychology was not prepared, he thought, for these referrals; it did not have a sufficient body of facts and principles upon which to base its answers. If it had been let alone to develop in its own way, as the physical and the biological sciences were during their infancy, it would in time have acquired this body of knowledge; but the demands from outside, which psychology sought to meet, brought hybrid concepts and problems and shoddy methods of study into psychology—concepts, problems and methods that were foreign to its development—hence the hodgepodge that psychology is today.

In his system, which he called dynasomatic psychology, Bentley sought to reëstablish psychology as an independent science and to consolidate and redintegrate its facts and principles, both old and new. His book was written with his usual circumspection. It was mimeographed and used for several terms as a text in the elementary courses at Cornell. During this period, Bentley held weekly seminars with the members of the staff using the book, sessions at which it was read and every sentence and paragraph combed through for clarity, style, and factual slips. Here, in this book, is Bentley at his best.

The New Field, published in 1934, is a functional psychology, resembling Brentano's act psychology more closely than American functionalism. It is the most logically consistent and comprehensive system presented since Brentano's *Psychologie vom empirischen Standpunkt.* Given his premises, his treatment follows along logically from beginning to the end, providing sanction for the entire field of psychology—including the normal adult, the fields of applied, child, comparative, developmental, and mental defects and disorders, many areas which other systems are unable to embrace.

The book had, however, little impact upon psychology, because Bentley was, first and foremost, a scholar and not a promoter. He had little ability and no inclination along promotional lines, and his students were too immature psychologically to assume that role for him. Psychology had, moreover, just passed through a long period of systematic promotion and it

[34] *Loc cit.*

was surfeited and cold toward its resumption.[35] The time was not ripe, therefore, for the introduction of a new system and before times again became favorable and his students had reached the stature when they might have acted as promoters, the feverish and traumatic years of the War were upon us. Since the War, psychology has shown little temper or ability to resume its systematic considerations.

Shortly after the War, about 1948, Bentley turned to what was to be his *magnum opus*. The central theme of this book was, as he confided to the writer, *"Whence, where, whither psychology*—not a title, but a guide; not another history of psychology, but a selection and an inferential use of men, writings, events, methods, external pressures, and the like."[36] He applied himself assiduously and unsparingly—almost frantically—to this self-appointed task. It was as if he realized that its accomplishment was to be a race with time. That he might devote all of his time and energy to this book, he withdrew from all other activities, social as well as professional—even to resigning from his editorship of this JOURNAL to which he was strongly attached by sentiment.

The thoroughness with which he prepared himself to write this book, which he entitled *Heritage and Attainment*, is shown in a letter received from him by the writer shortly before his last illness. He wrote: "I have now finished 400 yr. of physiological influence—from Jean Fernel's *Psychologie* of 1554 to the present. I have pared off all the 'fat' and non-essentials but the monster threatens to consume me. One item that has been richly rewarding has been the re-reading of the twenty-odd volumes of the Meckel-Müller's *Archiv* and Müller's *Handbuch* and *Handwörterbuch.*" After a life-time spent in general preparation and six years spent in special preparation, he was ready to write his book! It was, however, later than he thought; the time was not vouchsafed him, so this effort, like many of his other efforts, was lost to psychology.

It is a pity that Bentley's critical acumen, one of his outstanding characteristics, should have been brought to bear so frequently upon his own work. Had he been less self-critical, we should have shared a larger portion of his erudition.[37]

(b) Editorial. His erudition and critical acumen, which curtailed the publication of his books, well qualified him, on the other hand, for the

---

[35] See *Psychologies of 1925, op. cit.,* 1-412; *Psychologies of 1930,* Clark University Press, 1930, 1-497.

[36] Personal correspondence.

[37] It is also a pity, it might be added, that many other writers, past and present, have not had a modicum of Bentley's self-criticism. Had they had, we should then have been spared many of the badly conceived, hastily written books and articles that now clutter our shelves.

many editorial tasks that were crowded upon him and to which he gave, through most of his professional life, a large share of his time and energy. He was an editor of this JOURNAL for forty-eight years (coöperating editor, 1903-1926; co-editor, 1926-1950), of the *Psychological Index* for ten years (1915-1924), of the *Journal of Experimental Psychology* for four years (1926-1929); and an associate editor of the *Journal of Comparative Psychology* for fifteen years (1921-1935). In addition to these journals, he edited four monographic series of studies (three from the University of Illinois and one from Cornell) and was co-editor of two books.[38]

Bentley was an editor of the 'old school.' He did not merely mark copy for type size and correct it for spelling, punctuation, and the occasional split infinitive, but he also edited it for style. Articles he passed for publication had to be stylistically as well as factually correct. One of his special aversions was the adjectival use of nouns—an illiteracy imported into scientific writing during World War II—which spread so rapidly in psychology that he was moved to publish a Note in protest.[39]

Bentley edited as deeply as a good, clear style demanded. For young and inexperienced authors, whose manuscripts showed a glimmering of an idea, he frequently rewrote their articles entirely, returning his revisions for the authors' approval, study and profit. Whether the author was old or young, however, no manuscript ever passed under his editorial pen without being the better for it—as even Titchener attested. The debt that psychology and various psychologists owe Bentley for his editorial services is great.

(c) Experimental researches. As is usually true of most men who have become involved in administration, editing, and the writing of books, Bentley's chief experimental researches were performed in his youth and latterly only through his students. During Bentley's first period of service at Cornell, Titchener directed the researches of the graduate students, hence Bentley had time and leisure to conduct his own experiments. Among the numerous studies published by him, three played an important role in psychology: (1) "The memory image and its qualitative fidelity," which, as observed above, eventually led systematic psychology to abandon the image as a mental element; (2) "The synthetic experiment," which immediately became a paradigm of scientific procedure and was carried from research to research and laboratory to laboratory for the next quarter of a century; and (3) a study with L. M. Day on learning in the Paramecium, the interpretation of the results of which was the center of controversy for the following twenty

---

[38] For references to these books, see accompanying bibliography.
[39] Bentley, A noun modifies a noun modifies a noun modifies a noun, this JOURNAL, 57, 1944, 268 f.

years. It falls to the lot of few men to produce as many researches that have been so influential.

Bentley excelled as a teacher of seminars and of advanced courses; he seldom essayed to teach the elementary course. He had little patience with mediocrity—in attempting to fit square pegs in round holes—and was apt to be short and caustic with students who were "taking the course just for the credit" and not for the knowledge to be gained. To a student showing a real interest and a desire to know, he was kindly and patient and gave unsparingly of his time, though the student be at the foot of the class. He demanded high standards of endeavor and was ever ready to lend assistance to any one willing to work, but woe betide the negligent and indifferent student who by chance registered for one of his classes. He was especially intolerant of muddled thinking and the careless use of language, often indicating for students he deemed worthy and promising how papers and articles should be prepared by rewriting them himself.

His own style, though grammatically perfect, was difficult at best. It was complicated by his phraseology and use of words, by the polish he gave his compositions in their repeated revisions, by the number of examples given to illustrate a point, and by the allusions and figures of speech with which he enriched his manuscripts. These very excellences frequently attract so much attention to themselves that the reader finds that he must go back and re-read to pick up the meaning as a whole. Bentley was, as we have seen, enamored in his youth of Browning. It is possible and indeed probable that there is something of Browning in Bentley's style.

Bentley's death closed a long, full life, rich in accomplishments and services to the science to which he had dedicated himself. During the sixty-year span of his professional life (from his entrance into Cornell University in 1895 until his death in Palo Alto in 1955), he received all the honors within the power of his confrères to bestow upon him save the one that is usually accorded outstanding men of science. Why this exception, it is difficult for outsiders to understand; but better it be asked of one, why one was not elected than why one was. The answer, despite his outstanding record as an experimental scientist, teacher, author, and editor, may lie in the fact that he was *par excellence* a critic. All of his numerous book reviews and over half of the 150-odd articles from his pen are fearlessly critical. He was bold to criticize wherever criticism was pertinent; he let the chips fall where they would.

The University of Texas   KARL M. DALLENBACH

A BIBLIOGRAPHY OF THE WRITINGS OF
MADISON BENTLEY

By M. C. McGRADE and KARL M. DALLENBACH, University of Texas

I. BOOKS

(1) The Field of Psychology: A Survey of Experience, Individual, Social and and Genetic. New York, Appleton, 1924, xvi-545.

(2) The New Field of Psychology: The Psychological Functions and Their Government. New York, Appleton-Century, 1934, xxi-439.

II. EDITORIAL

(1) Coöperating editor, *The American Journal of Psychology,* from 1903 (Vol. 14) to 1925 (Vol. 36).

(2) Editor, *The Psychological Index,* from 1916 (Vol. 22) to 1925 (Vol. 31).

(3) Editor, Studies in Social and General Psychology from the University of Illinois, *Psychol. Monogr.,* 21, 1916 (No. 92), 1-115.

(4) Associate editor, *The Journal of Comparative Psychology,* from 1921 (Vol. 1) to 1935 (Vol. 20).

(5) Editor, Critical and Experimental Studies in Psychology from the University of Illinois, *Psychol. Monogr.,* 30, 1921 (No. 136), 1-94.

(6) Editor, *The Journal of Experimental Psychology,* from 1926 (Vol. 9) to 1929 (Vol. 12).

(7) Co-Editor, *The American Journal of Psychology,* from 1926 (Vol. 37) to 1950 (Vol. 63).

(8) Editor, Studies in Psychology from the University of Illinois, *Psychol. Monogr.,* 35, 1926 (No. 163), 1-151.

(9) Co-Editor (with K. M. Dallenbach and E. G. Boring), Washburn Commemorative Volume, *Amer. J. Psychol.,* 39, 1927, viii, 442.

(10) Co-Editor (with E. V. Cowdry), *The Problem of Mental Disorder.* New York; McGraw-Hill, 1934, x, 388.

(11) Editor, *Cornell Studies in Dynasomatic Psychology,* 1938, 255.

III. ARTICLES, NOTES, DISCUSSIONS

(1) The psychology of 'the grammar of science,' *Phil. Rev.,* 6, 1897, 521-526.

(2) The memory image and its qualitative fidelity, *Amer. J. Psychol.,* 11, 1899, 1-48.

(3) Current discussions of psychology and education, *Phil. Rev.,* 8, 1899, 104-108.

(4) The synthetic experiment, *Amer. J. Psychol.,* 11, 1900, 405-425.

(5) Illusions: I. The psychology of deception, *The Chicago Record-Herald,* June 24, 1901.

(6) Illusions: II. Geometry and the eye, *The Chicago Record-Herald,* July 1, 1901.

(7) Illusions: III. Geometrical illusions, *The Chicago Record-Herald,* July 8, 1901.

(8) Illusions: IV. Explaining geometrical illusions, *The Chicago Record-Herald,* July 15, 1901.

(9) Illusions: V. Illusions of the Parthenon, *The Chicago Record-Herald,* July 22, 1901.

(10) Illusions: VI. Roman and mediaeval architecture, *The Chicago Record-Herald,* July 29, 1901.

(11) Illusions: VII. Illusions in modern architecture, *The Chicago Record-Herald,* August 5, 1901.

(12) Illusions: VIII. Illusions at the Pan-American Exposition, *The Chicago Record-Herald,* August 12, 1901.

(13) Illusions: IX. Illusions at the Pan-American Exposition, *The Chicago Record-Herald,* August 19, 1901.

(14) Illusions: X. Illusions and interiors, *The Chicago Record-Herald,* August 26, 1901.

(15) Illusions: XI. Illusions in dress, *The Chicago Record-Herald,* September 2, 1901.

(16) Illusions: XII. Illusions in dress, *The Chicago Record-Herald,* September 9, 1901.

(17) Illusions: XIII. The psychology of magic, *The Chicago Record-Herald,* September 16, 1901.

(18) Illusions: XIV. The witchery of suggestion, *The Chicago Record-Herald,* September 23, 1901.

(19) Illusions: XV. Mental architecture, *The Chicago Record-Herald,* September 30, 1901.

(20) Illusions: XVI. Hypnotism and illusion, *The Chicago Record-Herald,* October 7, 1901.

(21) Illusions: XVII. The passing of superstition, *The Chicago Record-Herald,* October 14, 1901.

(22) The psychology of mental arrangement, *Amer. J. Psychol.,* 13, 1902, 269-293.

(23) President Minot on 'The Problem of consciousness in its biological aspects,' *Sci.,* N. S., 16, 1902, 386-391.

(24) The simplicity of color tones, *Amer. J. Psychol.,* 14, 1903, 92-95.

(25) Professor Calkins on mental arrangements, *Amer. J. Psychol.,* 14, 1903, 113-114.

(26) A critique of 'fusion,' *Amer. J. Psychol.,* 14, 1903, 324-336.

(27) (With E. B. Titchener) Ebbinghaus' explanation of beats, *Amer. J. Psychol.,* 15, 1904, 62-71.

(28) Professor Cattell's statistics of American psychologists, *Amer. J. Psychol.,* 15, 1904, 102-103.

(29) The psychological meaning of clearness, *Mind,* N. S., 13, 1904, 242-243.

(30) Standards of audition, *Sci.,* N. S., 19, 1904, 959-961.

(31) (With G. H. Sabine) A study in tonal analysis—I, *Amer. J. Psychol.,* 16, 1905, 484-498.

(32) The psychology of organic movements, *Amer. J. Psychol.,* 17, 1906, 293-305.

(33) (With M. F. Washburn) The establishment of an association involving color-discrimination in the creek chub, *Somotilus atromaculatus, J. comp. neur. Psychol.,* 16, 1906, 113-125.

(34) An observation table, *Amer. J. Psychol.,* 20, 1909, 278-279.

(35) Mental inheritance, *Pop. Sci. Mo.,* 75, 1909, 458-468.

(36) The historians, in Theodore Stanton (ed.), *Collection of British and American Authors,* New York, G. P. Putnam's Sons, 1909, 89-115; Leipzig, Bernhardt Tauchnitz, 1909, 107-135.

(37) (With S. E. Barnholt) Thermal intensity and the area of stimulus, *Amer. J. Psychol.,* 22, 1911, 325-332.

(38) (With L. M. Day) A note on learning in Paramecium, *J. Animal Behav.,* 1, 1911, 67-73.

(39) Sensation: General, *Psychol. Bull.,* 8, 1911, 79-80.

(40) (With E. G. Boring and C. A. Ruckmich) New apparatus for acoustical experiments, *Amer. J. Psychol.,* 23, 1912, 509-516.

(41) Sensation: General, *Psychol. Bull.,* 9, 1912, 97-98.

(42) Sensation: General, *Psychol. Bull.,* 10, 1913, 85-87.

(43) The study of dreams; a method adapted to the seminary, *Amer. J. Psychol.,* 26, 1915, 196-210.

(44) Sensation: General, *Psychol. Bull.,* 12, 1915, 100-102.

(45) Sensation: General, *Psychol. Bull.,* 13, 1916, 120-122.

(46) Prefatory note, *Psychol. Monogr.,* 21, 1916 (No. 92), iii.

(47) A preface to social psychology: University of Illinois Studies, *Psychol. Monogr.,* 21, 1916 (No. 92), 1-25.

(48) The psychological antecedents of phrenology: University of Illinois Studies, *Psychol. Monogr.,* 21, 1916 (No. 92), 102-115.

(49) Sensation, *Psychol. Bull.,* 14, 1917, 81-82.

(50) A note on the relation of psychology to anthropology, *J. phil. Psychol.,* 17, 1920, 345-349.

(51) Dynamical principles in recent psychology, *Psychol. Monogr.,* 30, 1921 (No. 136), 1-16.

(52) Leading and legibility, *Psychol. Monogr.,* 30, 1921 (No. 136), 48-61.

(53) (With Annett Baron) The intensive summation of thermal sensations, *Psychol. Monogr.,* 30, 1921 (No. 136), 84-94.

(54) A note on the 'expression' of simple feelings, *Psychol. Rev.,* 30, 1923, 326-327.

(55) Applied science and science applied, *Sci.,* 58, 1923, 69-70.

(56) The psychologies called "structural": Historical derivation, *Psychologies of 1925; Powell Lectures in Psychological Theory,* Clark Univ. Press, 1926, 383-393; also in *Ped. Sem.,* 33, 1926, 293-303.

(57) The work of the structuralist, *Psychologies of 1925, Powell Lectures in Psychological Theory,* Clark Univ. Press, 1926, 395-404; also in *Ped. Sem.,* 33, 1926, 304-313.

(58) The psychological organism, *Psychologies of 1925, Powell Lectures in Psychological Theory,* Clark Univ. Press, 1926, 405-412; also in *Ped. Sem.,* 33, 1926, 314-321.

(59) (With K. M. Dallenbach, E. G. Boring, and M. F. Washburn) An announcement, *Amer. J. Psychol.,* 37, 1926, 1.

(60) Experimental psychology in the Orient, *Amer. J. Psychol.,* 37, 1926, 154.

(61) (With O. F. Weber) The relation of instruction to the psychosomatic functions, *Psychol. Monogr.,* 35, 1926 (No. 163), 1-15.

(62) Qualitative resemblance among odors, *Psychol. Monogr., 35,* 1926 (No. 163), 144-151.

(63) The major categories of psychology, *Psychol. Rev.,* 33, 1926, 71-105.

(64) Environment and context, *Amer. J. Psychol.,* 39, 1927, 54-61.

(65) Is 'emotion' more than a chapter heading?, in *Feeling and Emotions: The Wittenberg Symposium,* ed. by M. L. Reymert, Clark Univ. Press, 1928, 17-23.

(66) The general index of the first thirty volumes of The American Journal of Psychology, *Amer. J. Psychol.,* 40, 1928, 523-524.

(67) 'Observer' and 'subject,' *Amer. J. Psychol.,* 41, 1929, 682-683.

(68) Deutsche Gesellschaft für Psychologie, *Amer. J. Psychol.,* 41, 1929, 689.

(69) The psychologist's interest in learning, *Arch. Otolaryngol.,* 10, 1929, 282-295.

(70) (With F. W. Hodges, E. B. Passano, C. A. Peerenboom, H. C. Warren, and M. F. Washburn) Instructions in regard to preparation of manuscript, *Psychol. Bull.,* 26, 1929, 57-63.

(71) Record of mutilated speech and music, *Amer. J. Psychol.,* 42, 1930, 115.

(72) Lacquers and celluloids for colored surfaces, *Amer. J. Psychol.,* 1930, 116.

(73) First International Congress on Mental Hygiene, *Amer. J. Psychol.,* 42, 1930, 147.

(74) Titchener's *Systematic Psychology: Prologomena, Amer. J. Psychol.,* 42, 1930, 148-153.

(75) Another note on the observer in psychology, *Amer. J. Psychol.,* 42, 1930, 320.

(76) A correction: Allport's 'social facilitation,' *Amer. J. Psychol.,* 1930, 320-321.

(77) (With C. C. Ring) The effect of training upon the rate of adult reading, *Amer. J. Psychol.,* 42, 1930, 429-430.

(78) (With R. H. Gundlach) The dependence of tonal attributes upon phase, *Amer. J. Psychol.,* 42, 1930, 519-543.

(79) (With W. H. Mikesell) Configuration and brightness contrast, *J. exper. Psychol.,* 13, 1930, 1-23.

(80) A psychology for psychologists, in *Psychologies of 1930,* Worcester, Clark Univ. Press, 1930, 95-154.

(81) The work of the Division of Anthropology and Psychology in the National Research Council, *Amer. J. Psychol.,* 43, 1931, 646-658.

(82) Research in psychology and its bearing upon the social sciences in *Essays on Research in the Social Sciences,* Washington, D.C., Brookings Institute, 1931, 123-135.

(83) In quest of glacial man: a plan for cooperation between excavators and the representatives of the sciences of man and of the earth, *Repr. & Circ. Ser. Nat. Res. Council,* 1931 (No. 100), 1-120.

(84) Psychology's family relations among the sciences, *Sci.,* 73, 1931, 113-117.

(85) The National Research Council's fund to support research, *Amer. J. Psychol.* 44, 1932, 195-196.

(86) Exchange of odd volumes and partial sets of psychological journals, *Amer. J. Psychol.,* 44, 1932, 196.

(87) (With E. J. Varon) An accessory study of 'phonetic symbolism,' *Amer. J. Psychol.,* 45, 1933, 76-86.

(88) Mind, body, and soul in medical psychology, *Amer. J. Psychol.,* 45, 1933, 577-591.

(89) Man: Dust or diety?, *Sci. Mo.,* 37, 1933, 365-368.

(90) Preface, *The Problem of Mental Disorder,* ed. by Madison Bentley and E. V. Cowdry, 1934, v-vii.

(91) The character of the problem, *The Problem of Mental Disorder,* Sec. I, 1934, 1-6.

(92) Current points of view, *The Problem of Mental Disorder,* Sec. II, 1934, 7-13.

(93) Summary of current points of view, *The Problem of Mental Disorder,* Sec. II, 1934, 91-95.

(94) The supporting sciences: Present contributions and future research, *The Problem of Mental Disorder,* Sec. III, 1934, 96-110.

(95) General and experimental psychology, *The Problem of Mental Disorder,* Sec. III, 1934, 275-308.

(96) Summary of present contributions and future research, *The Problem of Mental Disorder,* Sec. III, 1934, 372-376.

(97) Comments and reflections, *The Problem of Mental Disorder,* Sec. IV, 1934, 377-381.

(98) A remark on the new forms of the synchronous chronoscope, *Amer. J. Psychol.,* 47, 1935, 322.

(99) A new journal for the social sciences, *Amer. J. Psychol.,* 48, 1936, 174.

(100) (With E. G. Boring and M. F. Washburn) Editorial note, *Amer. J. Psychol.,* 48, 1936, 177.

(101) Conjunctive research in the sciences of life, *Amer. J. Psychol.,* 48, 1936, 512-519.

(102) The psychologist's uses of neurology, *Amer. J. Psychol.,* 49, 1937, 233-264.

(103) 'Hundreds have already adopted,' *Amer. J. Psychol.,* 49, 1937, 678-679.

(104) The nature and uses of experiment in psychology, *Amer. J. Psychol.,* 50, 1937, 452-469.

(105) The problems of mental disability in England, *Amer. J. Psychol.,* 51, 1938, 1-17.

(106) The psychological background of the doctoral candidate, *Amer. J. Psychol.,* 51, 1938, 170-172.

(107) Cornell studies in dynasomatic psychology, *Amer. J. Psychol.,* 51, 1938, 203-224.

(108) Retrospect and prospect, *Amer. J. Psychol.,* 51, 1938, 357-360.

(109) Oxygen-tension and 'the higher mental processes,' *Amer. J. Psychol.,* 52, 1939, 72-82.

(110) (With R. M. Lindner) A functional and dynasomatic study of emoving, *Amer. J. Psychol.,* 52, 1939, 186-209.

(111) Individual psychology and the psychological varieties, *Amer. J. Psychol.,* 52, 1939, 300-301.

(112) Monographs of the Society for Research in Child Development, *Amer. J. Psychol.,* 52, 1939, 477-478.

(113) (With E. G. Boring and K. M. Dallenbach) The editorial board, *Amer. J. Psychol.,* 53, 1940, 6.

(114) The 'best' psychologists, *Amer. J. Psychol.,* 54, 1941, 439.

(115) (With E. G. Boring and K. M Dallenbach) The department of apparatus notes, *Amer. J. Psychol.,* 55, 1942, 114.

(116) (With E. G. Boring and K. M Dallenbach) William James: 1842-1910, *Amer. J. Psychol.,* 55, 1942, 309.

(117) A general index to Volumes 31-50 of *The American Journal of Psychology, Amer. J. Psychol.,* 55, 1942, 582.

(118) (With E. G. Boring) Erratum, *Amer. J. Psychol.,* 55, 1942, 583.

(119) (With E. G. Boring) An announcement, *Amer. J. Psychol.,* 55, 1942, 583.

(120) Erratum, *Amer. J. Psychol.,* 56, 1943, 133.

(121) "I do admire an ardent soul," *Amer. J. Psychol.,* 56, 1943, 301-302.

(122) Where does thinking come in? *Amer. J. Psychol.,* 56, 1943, 354-380.

(123) Athena Nike, *Amer. J. Psychol.,* 56, 1943, 450.

(124) The theater of living in animal psychology, *Amer. J. Psychol.,* 57, 1944, 1-48.

(125) A psychological sketch of the young child, *Amer. J. Psychol.,* 57, 1944, 206-235.

(126) Tools and terms in recent researches: Is fatigue tiring?; The cat's theater of living; A noun modifies a noun modifies a noun, *Amer. J. Psychol.,* 57, 1944, 264-269.

(127) Tools and terms in recent researches: *Domi et foras, Amer. J. Psychol.,* 57, 1944, 424-426.

(128) Tools and terms in recent researches: Divide and subdivide, *Amer. J. Psychol.,* 57, 1944, 577-579.

(129) An editorial announcement, *Amer. J. Psychol.,* 57, 1944, 579-580.

(130) Erratum, *Amer. J. Psychol.,* 58, 1945, 117.

(131) Sanity and hazard in childhood, *Amer. J. Psychol.,* 58, 1945, 212-246.

(132) The reconstructing teens and the stabilizing twenties, *Amer. J. Psychol.,* 58, 1945, 324-360.

(133) Tools and terms in recent researches: Gizzard and psyche, *Amer. J. Psychol.,* 58, 1945, 113-117.

(134) Tools and terms in recent researches: Laboratory, field or ward, testing and reckoning rooms, library and gymnasium, *Amer. J. Psychol.,* 58, 1945, 394-397.

(135) Tools and terms in recent researches: The probe; the meter; the report, *Amer. J. Psychol.,* 59, 1946, 155-160.

(136) A graduate fellowship in experimental psychology, *Amer. J. Psychol.,* 59, 1946, 161.

(137) (With E. G. Boring) An editorial announcement, *Amer. J. Psychol.,* 59, 1946, 161.

(138) Life with and without institutional guidance, *Amer. J. Psychol.,* 59, 1946, 382-400.

(139) Tools and terms in recent researches: Stops and pauses, *Amer. J. Psychol.,* 59, 1946, 463-468.

(140) (With K. M. Dallenbach) An editorial announcement, *Amer. J. Psychol.,* 59, 1946, 495.

(141) Psychologizing for the multitude, *Amer. J. Psychol.,* 59, 1946, 722-726.

(142) A psychologist's reflection upon Howell's *Physiology, Amer. J. Psychol.,* 60, 1947, 420-432.

(143) Suggestions toward a psychological history of the hominids *Amer. J. Psychol.,* 60, 1947, 479-501.

(144) An acknowledgment, *Amer. J. Psychol.,* 60, 1947, 646.

(145) Advancement of understanding and advancement of professional service, *Amer. J. Psychol.,* 61, 1948, 111-118.

(146) The Harvard case for psychology, *Amer. J. Psychol.,* 61, 1948, 275-282.

(147) Factors and functions in human resources, *Amer. J. Psychol.,* 61, 1948, 286-291.

(148) The new edition of Sherrington's *Integrative Action, Amer. J. Psychol.,* 61, 1948, 562-573.

(149) Samiksa: Journal of the Indian Psycho-Analytical Society, Amer. J. Psychol., 61, 1948, 583.

(150) Scientific American, Amer. J. Psychol., 61, 1948, 583-584.

(151) Who is to bear primary responsibility for the psychological disorders?, *Amer. J. Psychol.,* 62, 1949, 257-265.

(152) An early 'fish target,' *Amer. J. Psychol.,* 63, 1950, 260-261.

(153) Annual review of psychology, *Amer. J. Psychol.,* 63, 1950, 289-290.

(154) Excerpta medica, *Amer. J. Psychol.,* 63, 1950, 290-291.

(155) A new Brazilian journal, *Amer. J. Psychol.,* 63, 1950, 291.

(156) Early and late metric uses of the term *distance, Amer. J. Psychol.,* 63, 1950, 619.

(157) Forecast, timing, and other primary factors in government of certain biochemical systems, *Amer. J. Psychol.,* 65, 1952, 329-345.

QUANTITATIVE DENOTATIONS OF COMMON TERMS AS A FUNCTION OF BACKGROUND

By Harry Helson and Robert S. Dworkin, University of Texas, and Walter C. Michels, Bryn Mawr College

Language has many words like 'few,' 'some,' and 'many,' that denote quantity more or less precisely. Words and phrases are often used as substitutes for numbers in estimating magnitudes of all kinds. The information conveyed by such terms varies with context and personal factors. Thus "Many died in the Cocoanut Grove disaster" refers to dozens, or hundreds at most, whereas "Many died in the Second Battle of the Marne" is a statement about hundreds of thousands. The number meant by the word 'many' obviously differs in the two contexts which must be known if the estimates are to be correctly interpreted.

Words with quantitative denotations are also employed to communicate evaluative judgments based largely on subjective or emotional states. If strong emotional biases are operative in forming judgments even quantitative terms may be greatly altered in meaning. Thus the society matron who asserts that 'nobody' attended a party given by a rival does not mean that not a single person attended or even that the number of guests was small, but that no one of social importance or from her set attended. Evidently the information and meanings conveyed by language depend upon more than the information conveyed by individual words, phrases, or even sentences.

The influence of context on information has been dealt with in a number of publications concerned chiefly with types of *stimuli* and contexts of *stimuli* without regard to their effects on the organism.[1] In this study we are interested in testing an hypothesis concerning the effect of background on the organism and, in turn, how this affects the response. According to adaptation-level (AL) theory,[2] the effect of stimuli depends upon background and residual factors as well as upon the stimuli them-

[1] G. A. Miller, What is information measurement? *Amer. Psychologist*, 8, 1953, 3-11.

[2] Harry Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this Journal, 60, 1947, 1-29; W. C. Michels and Harry Helson, A reformulation of the Fechner law in terms of adaptation-level applied to rating-scale data, *ibid.*, 62, 1949, 355-368.

selves. Variation in background should, therefore, cause a change in response. Identical stimuli when presented with different backgrounds will elicit different responses if our fundamental assumption is correct. Differences in responses among individuals exposed to the same stimulus-background conditions are attributable to differences in residuals since so-called subjective factors play a large role in language. In this study we shall be concerned mainly with background effects, leaving residual effects for a later report.

Two previous studies of verbal materials fit into the theoretical approach underlying the present investigation. Woodworth and Sells found three main sources of error in a study of syllogistic reasoning: (1) ambiguity of certain words like 'some'; (2) the atmosphere created by certain terms in the premisses, predisposing toward certain conclusions; and (3) the caution or wariness of Ss in drawing conclusions from universal and affirmative propositions.[3] From the point of view of adaptation-level theory these three sources of error can be regarded as springing from the material given (the stimulus-objects), from context or background, and from residual or individual (subjective) factors. All three may operate to produce one or another conclusion. An approach to the problem of subjective factors in word connotations formally similar to ours was made independently by Mosier in a psychometric study of the pleasantness-unpleasantness quality of 296 words such as 'normal,' 'disgusting,' 'excellent.'[4] He suggested that the meaning of a word may be quantitatively described as: $M = x + i + c$, where M is the meaning of the word, x is that part of the meaning which is constant from person to person and context to context, i is the part that varies from person to person, and c is the part that varies from context to context. On the basis that i and c often cannot be distinguished, Mosier simplified the formula to read: $M = x + v$, where v is a variable including individual and contextual variables.

In the present study the influence of background on the quantitative denotations of 26 common words and phrases is investigated. According to AL theory the effect of a stimulus is a function of the momentary state of adaptation of the organism and, if this can be specified in a given situation, responses should be predictable. The AL has been found to be a weighted geometric mean of three classes of stimuli: (1) stimuli immediately in the focus of attention and being responded to; (2) all other stimuli in the field or recently experienced which form the background of stimulation; and (3) effects of more remote stimulation called the residual stimulus.[5] Change in any of these three classes of stimuli may alter AL with the result that responses to stimuli will be modified. A high AL pre-

[3] R. S. Woodworth and S. B. Sells, An atmosphere effect in formal syllogistic reasoning, *J. exper. Psychol.*, 18, 1935, 451-460.

[4] C. I. Mosier, A psychometric study of meaning. *J. soc. Psychol.*, 13, 1941, 123-140.

[5] Michels and Helson, *op. cit.*

disposes to low or negative responses while low *AL* predisposes to high or positive responses. Concretely we may set up the following hypothesis to test in the present study.

Other things equal, the higher the background number against which quantitative terms are judged, the higher will be the resultant *AL* and the smaller will be the percentages denoted by the terms; conversely, the lower the background number, the lower will *AL* be and the higher will be the percentage denotations of the terms.

Procedure. Twenty-six commonly used words and phrases such as 'all,' 'almost all,' 'most,' 'quite a few,' 'practically nobody' were given to four groups of *S*s with instructions to indicate the numbers conveyed by each of the terms while keeping a given number in mind as background. Each group consisted of 75 men and women selected at random from classes in various courses at the University of Texas. The four background numbers were: 100; 1232; 144,690; and 1,728,583. The groups judging with these background numbers will be referred to as Group I, II, III, and IV, respectively. These numbers, with the exception of 100, were chosen as representative of fairly small to very large magnitudes and were also selected to prevent round number-judgments, or judgments in terms of percentages, or easily remembered responses. Group I served as the control group.

Instructions. Forms containing the 26 items were distributed to 300 *S*s with the following explanatory remarks given orally.

In ordinary conversation people frequently substitute descriptive terms for numbers to convey the concept of size or amount. It is the purpose of this study to discover what quantitative meanings some commonly used terms have. Kindly indicate how many people are designated by the following descriptive phrases by putting numbers in the spaces provided. Be sure to give your first impression. Do not try to be too exact as we wish the type of reaction which you would give in conversation or in ordinary verbal usage.

To provide the background number, each form contained the following statement:

Given ———— people, how many people would you include in the following cases?

One of the background numbers (1232, 144,690, or 1,728,583) was inserted in the blank space. In the case of Group I, the control group, the instructions were:

Indicate the percentage of people (or number in 100) designated by the following terms.

After the tenth item, the *S*s were reminded of the background number by the sentence:

Remember, the number of people given is ————.

Treatment of data. Means and *SD*s were computed for each word or phrase for each group (or background number). In the case of the three groups whose background numbers were larger than 100 the means and *SD*s were divided by the background number to yield what will be referred to as *derived percentages* to

make them comparable with each other and with the estimates of the control group. Comparisons of the differences between the derived percentages and between these and the control percentage should show the influence of the four backgrounds. All differences between means were tested for significance by means of the *t*-test for paired replicates since the same terms were used under all conditions.

Results: (a) Background factors. The mean estimates and the SDs of the distributions of estimates made by 75 Ss in each of of four groups

TABLE I

DERIVED MEANS AND SDs OF 26 TERMS

Background numbers

Terms	100		1,232		144,690		1,728,583	
	M	SD	M	SD	M	SD	M	SD
1. All	98	3	97	7	97	11	97	7
2. Everybody	95	7	96	6	96	21	97	12
3. Practically everybody	89	9	87	17	90	17	84	20
4. Almost all	88	12	88	10	89	16	85	20
5. Almost everybody	87	8	75	33	85	20	84	17
6. Most	78	12	72	28	78	20	77	16
7. Generally	71	13	74	25	71	24	73	26
8. A majority	64	11	59	25	59	19	60	18
9. Too many	62	25	75	23	57	33	62	34
10. A lot	58	18	65	20	49	25	47	26
11. Many	58	20	62	14	50	35	45	28
12. A considerable number	54	22	63	14	52	25	46	24
13. Fairly common	53	19	54	30	54	26	48	26
14. Quite a few	43	19	45	23	38	26	33	26
15. A minority	31	11	30	18	20	19	27	21
16. Too few	21	15	17	16	19	22	15	20
17. Some	18	10	20	19	12	18	8	14
18. Several	14	14	9	14	5	15	4	13
19. Few	12	7	11	8	4	11	4	8
20. Uncommon	11	10	10	12	5	11	5	12
21. Hardly anybody	8	10	11	15	3	4	2	4
22. Scarcely anybody	6	4	6	10	4	9	2	6
23. Practically nobody	6	6	10	17	2	4	6	19
24. Almost no one	4	3	6	4	3	4	3	14
25. Nobody	2	3	0*	5	1	2	0*	1
26. None	0*	1	0*	3	0*	2	0*	0*
Grand Mean	43.5		43.8		40.1		39.0	

* Less than 1.

judging 26 terms (words or phrases) with one of four numbers serving as background are given in Table I. Do the background numbers exert a systematic influence upon the quantitative denotations of the items? According to the hypothesis the background should have a definite and regular effect on the judgment of the 26 terms. From the data in Table I it is seen that the mean estimates of the items decrease as the background num-

bers increase, with the exception of the 1232 background. The grand mean falls as the background number increases in accordance with the hypothesis as shown by grand means of 43.5 for 100 background, 43.8 for 1232 background, 40.1 for 144,690 background, and 39.0 for 1,728, 583 background. The differences between the 100 and 1232 backgrounds and between the 144,690 and 1,728,583 backgrounds are not statistically significant; but the differences between the two lower backgrounds and the two higher backgrounds are highly significant as shown by the *t*-values in Table II. These findings point to a levelling effect at the low and high

TABLE II

SIGNIFICANCE OF DIFFERENCES BETWEEN GRAND MEANS OF
CONTROL AND DERIVED PERCENTAGES

(Below the difference between means is the *t*-value. Starred *t*s are statistically
significant; at 1% level, $t=2.79$.)

Background numbers:	1,232	144,690	1,728,583
100	0.33	3.40	4.51
	0.34	4.85**	5.60**
1232		3.73	4.84
		2.93**	3.66**
144,690			1.11
			1.68

ends of the background continuum. Apparently 100 and 1232 are lower bounds while 144,690 and 1,728,583 are upper bounds so far as the influence of background on the estimates is concerned. These results gain added significance when we consider that they come from four independent groups of *S*s. We may consider our first hypothesis verified within the upper and lower bounds noted: the higher the background number against which quantitative terms are judged, the lower will be the relative numbers denoted by the terms.

Inspection of the means and *SD*s of the items in Table I shows that there is considerable variation from *S* to *S* in the quantitative denotation of most of the words and phrases. Thus 'all' may denote from 90 to 100% of the background number while 'nobody' and 'none' may range from 0 to around 4%. While some of the variation may be due to particular background numbers, *e.g.* the difficulty of estimating exactly 51% of 1,728,583 for 'majority,' this does not explain the variance of such words as 'nobody' and 'none' because it would have been possible to report zero with all backgrounds if these terms meant this to all *S*s. Standard deviations of the estimates show there was more or less scatter in the quantities

denoted by all the words and phrases. Thus 'generally' may denote as little as 50% of the background number or as much as 99%, depending upon the individual estimator. The wide range of values reported for 'majority' shows that many people do not stick by the dictionary meaning of the term (51%) since the estimates vary between 40 and 83% of the background numbers. If the percentages are translated into the original absolute numbers reported, the variations from S to S, as well as from background to background, become even more impressive.

The terms in Table I may be divided into three groups on the basis of the percentages reported with all backgrounds. The first group, consisting of the first seven terms, yields estimates above 70%; the second group, consisting of the next eight terms, includes estimates between 25 and 60%; and the third group, containing the last 11 terms, contains estimates from zero to 21%. The average change of the first group from the two smaller background numbers to the two higher background numbers is 2.0% which is not statistically significant as shown by the t-tests in Table III. The absolute numbers denoted by the terms in this group must therefore have risen with background number to preserve almost perfect constancy percentagewise.

Terms in the second group, Items 8-15 in Table I, drop from an average of 55 to 47% as background numbers are raised from 100 and 1232 to 144,690 and 1,728,583. The t-tests in Table III show the differences between the two lower and two higher backgrounds to be highly significant. Similarly the terms in group three, Items 16-26, drop from an average of 9.2 with the two lower backgrounds to 4.9% with the two higher backgrounds and this difference also is highly significant.

The division of the 26 terms into the three groups just discussed is based upon the magnitudes denoted by the terms, hence upon their quantitative meanings. As a result each of the groups contains terms that, on another basis, would go elsewhere. If we take as the basis of grouping lack of change percentagewise from background to background, the last three terms in Group III, 'almost no one,' 'nobody,' and 'none,' should go into Group I since the terms in this group do not change significantly with background number. These terms are exact opposites in meaning to 'almost everybody,' 'everybody,' and 'all' which do not change percentagewise with background. In the theoretical analysis which follows the terms have been regrouped on the basis of variability rather than meaning. While the three new groups parallel the ones just discussed, in that the first contains items that do not change from background to background and the other two groups contain items that shift significantly percentage-

wise, the new basis of grouping lends itself to better theoretical analysis. We therefore turn now to the theoretical treatment.

(b) Theoretical. The fact that the words and phrases studied here show both stability of meaning and shifts of meaning as the frame of reference changes, just as psychophysical judgments do, suggests that it may

TABLE III

SIGNIFICANCE OF DIFFERENCES BETWEEN DERIVED ESTIMATES OF THREE SUB-GROUPS
OF TERMS JUDGED ON THREE BACKGROUNDS
(Starred *t*-values are statistically significant.)

	Background numbers:	1,232	144,690	1,728,583
Items 1-7	100	2.43 1.27	0.00 0.00	1.29 1.30
	1,232		2.43 1.48	1.14 0.69
	144,690			1.29 0.38
Items 8-15	100	−3.75 1.87	5.75 4.22**	6.62 4.44**
	1,232		9.25 3.94**	10.37 3.93**
	144,690			1.12 0.64
Items 16-26	100	0.22 0.79	4.04 4.63**	4.86 4.42**
	1,232		3.82 3.41**	4.64 4.23**
	144,690			0.82 0.81

be possible to apply the quantitative theory used for the interpretation of psychophysical data to semantic problems. On the basis of a few reasonable assumptions, it has been demonstrated that a simple relation connects the judgment J of a physical stimulus with the physical intensity X of the stimulus.[6] The functional dependence is given by:

$$J = N + K \log_e (X/A), \ldots\ldots [1]$$

where A is the adaptation-level, a weighted geometric mean of all stimuli, past and present, which affect the judgment, where the judgment is expressed as the ordinal number of a category in a properly chosen rating scale containing $2N + 1$ categories numbered from zero to $2N$, and

---

[6] Michels and Helson, *op. cit.*

where K is a calculable function of N. Equation [1] has been tested under a variety of conditions and has been verified experimentally in every case to which it has been applied.

It has been shown that Equation [1] applies equally well to absolute and comparative judgments.[7] In the former case, the adaptation-level is strongly influenced by the residual stimulus, the magnitude of which is determined by the physiology of the organism and by the totality of experiences which the abserver has had previous to the time at which a judgment, or set of judgments, is made. The effect of this residual stimulus can be detected whenever the rating scale employs terms such as 'heavy,' 'medium,' and 'light,'—terms which carry absolute meaning. On the other hand, if the scale employs relative terms, such as 'heavier,' 'equal,' and 'lighter,' A is the comparative adaptation-level and is influenced most strongly by the standard in terms of which the comparisons are made. In either case, A is the magnitude of the stimulus which evokes the central or neutral judgment, 'medium' or 'equal.'

As was stated above, K is a function of N, and, therefore, of the number of categories in the rating-scale. When N is large, *i.e.* when fine discriminations are being attempted, and when the logarithms are to the base e, K becomes equal to N to a high degree of approximation. We may, therefore, rewrite Equation [1] as $J/N = 1 + \log_e X - \log_e A$. When the judgments are expressed on a percentage scale, and when the conditions are such that judgments greater than 100% do not occur, we can express the percentage judgment, J', as:

$$J' = J/2N = \frac{1}{2} + \frac{1}{2} \log_e X - \frac{1}{2} \log_e A. \ldots \ldots [2]$$

Hence the 50% judgment (0.50) will occur for $X = A$.

Most of the words and phrases used in this study fall into two distinct classes. Some of them, such as 'all,' 'a majority,' 'a minority,' clearly derive their meanings from the background number and express *relative* magnitudes. These we shall call Class R words and phrases. Others, such as 'many,' 'several,' 'few,' carry some *absolute* meaning, irrespective of the background number, although not as much as do numerical terms, such as 2, 10, 100. These we shall group into Class A. Class R words may be expected to hold constant their relative meanings with respect to the background number, since the adaptation-level may be expected to shift proportionately with that number. On the other hand, the adaptation-level

---

[7] Harry Helson, W. C. Michels, and Artie Sturgeon, The use of comparative rating scales for the evaluation of psychophysical data, this JOURNAL, 67, 1954, 321-326.

for Class A words may be expected to involve a constant residual factor as well as a factor determined by the background number, and the meaning of the terms expressed as percentages, may be expected to change as the background number changes.

Most of the words and phrases can be separated into the two classes by inspection. A better means of categorization can be obtained if we place in

TABLE IV

CLASSIFICATION OF TERMS ON THE BASIS OF BACKGROUND NUMBER

(Percentage changes are differences of the mean derived percentages for the two large background numbers and those for the two small background numbers.)

Term	Percentage change (D)	SD of change ($\sigma_{diff.}$)	$D/\sigma_{diff.}$	Class
1. All	−0.5	0.9	0.6	R
2. Everybody	+1.0	1.7	0.6	R
3. Practically everybody	−1.0	1.9	0.5	R
4. Almost all	−1.0	1.7	0.6	R
5. Almost everybody	+3.5	2.5	1.4	R
6. Most	+2.5	2.2	1.1	R
7. Generally	−0.5	2.6	0.2	R
8. A majority	−2.0	2.2	0.9	R
9. Too many	−9.0	3.4	2.6	AR
10. A lot	−13.5	2.6	5.2	A
11. Many	−12.5	3.0	4.2	A
12. A considerable number	−9.5	2.5	3.8	AR
13. Fairly common	−2.5	3.0	0.8	R
14. Quite a few	−8.5	3.0	2.8	AR
15. A minority	−7.0	2.0	3.5	AR
16. Too few	−2.0	2.1	1.0	R
17. Some	−9.0	1.8	5.0	A
18. Several	−7.0	1.6	4.4	A
19. Few	−7.5	1.0	7.5	A
20. Uncommon	−5.5	1.4	4.0	AR
21. Hardly anybody	−7.0	1.1	6.4	A
22. Scarcely anybody	−3.0	0.9	3.3	AR
23. Practically nobody	−4.0	1.6	2.5	AR
24. Almost no one	−2.0	0.9	2.2	AR
25. Nobody	−0.5	0.4	1.2	R
26. None	0.0	0.2	0.0	R

Class A those words which show a highly significant change of their derived means as the background number changes, and in Class R those words which do not show significant change. As a measure of the change we may take the difference between the average derived mean for the two small background numbers and that for the two large background numbers. This grouping of 150 estimates into each mean improves the averaging and is justified by the results shown in Table III. As a criterion of the significance of the change, we may take the ratio of the absolute value of the difference to the standard deviation of the difference. Table IV

shows the difference (D), the standard deviation of the difference ($\sigma_{diff.}$), and the critical ratio for each term tested.

The choice of the value of ($D/\sigma_{diff.}$) at which we shall make the break between classes must be somewhat arbitrary. The classification shown in the last column of Table IV, is obtained by saying that values of ($D/\sigma_{diff.}$) greater than 4.0 indicate highly significant changes, and that values less than 2.0 do not. Terms which exhibit ratios between these limits have been placed in an intermediate Class AR. The groupings of Class A and Class R terms are certainly close to those that might have been ex-

TABLE V

DIFFERENCES IN THE DERIVED MEANS FOR SUCCESSIVE CLASS A TERMS

Terms	Background number				$M_{diff.}$
	100	1,232	144,690	1,728,583	
10. A lot					
11. Many	0	3	−1	2	1
17. Some	40	42	38	37	39
18. Several	4	11	7	4	6
19. Few	2	−2	1	0	0
21. Hardly anybody	4	0	1	2	2

pected, while several of the words and phrases in Class AR would have presented difficulties in placement on an a priori basis.

The terms of Class A are those of greatest interest, and most of the following discussion will be devoted to them. If we suppose that the percentage judgments of these words and phrases is represented by Equation [2], with the adaptation-level being shifted by the changes in the background number, we should find that each term can be assigned a 'strength' X, which is invariant. To test this, we may take the differences between the derived means for the Class A terms for each background number. These differences are shown in Table V. Inspection of this table shows that, as anticipated, the difference in the meanings of any two words in Class A changes little with background number, although the derived means themselves change appreciably. This is entirely in accord with Equation [2] and encourages us to attempt further application of that equation.

Another fact which emerges from an examination of Tables IV and V is that 'a lot' and 'many' are essentially synonymous, as are 'several'

and 'a few.' We may, therefore, reduce the list of independent Class *A* terms from six to four, and take the average of the derived means for the pairs of synonyms. The derived means and their average differences, computed on this basis, are shown in Table VI.

With the general behavior of Equation [2] verified, we may now use

TABLE VI

REDUCED SUMMARY OF DERIVED MEANS FOR CLASS A TERMS

Term	Background number				Average change in differences between terms
	100	1,232	144,690	1,728,583	
10, 11. Many (A lot)	58	64	50	46	
					40
17. Some	[18	20	12	8	
					7
18, 19. Several (Few)	13	10	4	4	
					2
21. Hardly anybody	8	11	3	2	

it in combination with our data to determine the strength of each of the terms and the way in which the adaptation-level shifts with the background number. The differences in $\log_e X$ are, of course, 1/50 times the numbers in the last column of Table V. Since these specify only ratios of strengths, we are free to choose the strength of any one of the Class *A* terms arbitrarily. For present purposes, we shall choose the strength of

TABLE VII

STRENGTH OF CLASS A TERMS

Term	$\log_e X$	Strength (X)
10, 11. Many (A lot)	0.80	2.23
17. Some	0.00	1.00
18, 19. Several (Few)	−0.14	0.87
21. Hardly anybody	−0.18	0.84

'some' as unity. Then, from Equation [2], we may find the strengths of the various terms as stated in Table VII.

As was stated above, the adaptation-level *A* is a weighted geometric mean of all stimuli which affect the judgment. The two factors which are apparent in the present case are the background number, which we shall denote by Σ and the residual stimulus, which is determined by the past experience of the person using the term. This residual undoubtedly varies somewhat from person to person; we shall denote its average value

by R. It and the appropriate weighting factors must be determined from the experimental data. We write:

$$\log_e A = r \log_e R + (1 - r) \log_e \Sigma. \ldots \ldots \ldots [3]$$

A best average fit to the experimental values of Table VII is given by the values: $r = 0.98$, $\log_e R = 0.504$, $R = 1.66$. Fig. 1 shows the four
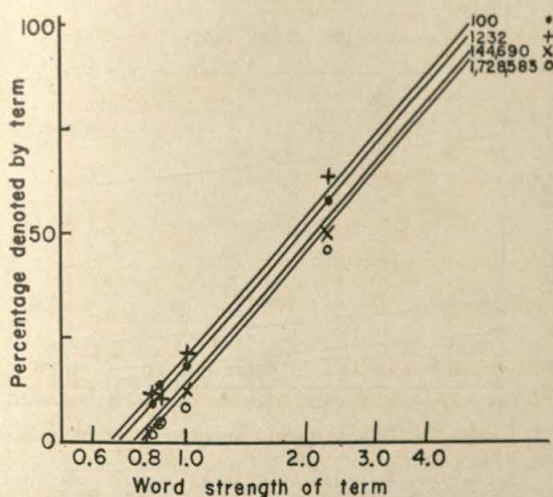
FIG. 1. TEST OF THE APPLICATION OF THE REFORMULATED FECHNER LAW
The word strengths plotted as abscissae are those determined from Table VII. The straight lines are plotted in accordance with the equation:
$$J' = \tfrac{1}{2} + \tfrac{1}{2} \log_e X - \tfrac{1}{2} (0.98 \log_e R + 0.02 \log_e \Sigma).$$

plots of percentage meaning against word strength, which result when these values are used in Equations [3] and [2]. While the fit to the experimental points shown is not perfect, the Equations do indicate the general trend of Class A words satisfactorily, particularly in view of the large standard deviations of the distributions of meanings.

A somewhat clearer presentation of this analysis is given by the plots shown in Figs. 2 and 3. In the former of these, the percentage meaning of the four independent Class A words is plotted as a function of the background number, according to Equations [2] and [3]. Fig. 3 shows a plot of the number represented by each term as a function of the same variable. In both cases the experimental points are shown and plots of the meaning of one Class R term is shown by the dashed lines, for comparison purposes.

Since the Class R terms do not shift in meaning with background number, there is no way of checking their behavior against Equation [2], as was done in Table V for Class A terms. If we assume that the test of Class A terms justifies the application of the treatment to Class R terms,
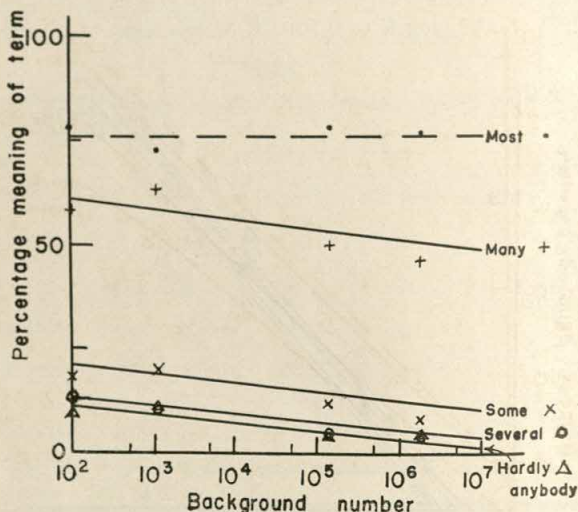
FIG. 2. PERCENTAGES DENOTED BY CLASS A TERMS AS A FUNCTION OF BACKGROUND NUMBER

The keying symbols used to plot the observed meanings are shown at the right. The term 'most' is included as an example of a Class R term whose percentage meaning is independent of background. The straight lines are computed on the basis of Equations [2] and [3].

we may use the differences in meaning of the latter to define a set of term strengths. These strengths, together with the synonyms, are shown in Table VIII, which has been constructed by choosing the strength of a (non-existent) term with an average meaning of 25% as unity. Strictly speaking, of course, the strengths of Class A and Class R terms cannot be compared, but this choice of unit strength for Class R terms insures that Equations [2] and [3] will work for both classes, with the same choice of the residual R. The only difference is that $r = 0.98$ for Class A and $r = 1.00$ for Class R.

Summary and Conclusions. We have found the hypothesis with which we began this study is supported by the data: as the reference number increases the percentage meaning of common words and phrases denoting quantity decreases significantly in the case of over half the words and

phrases. The remaining words and phrases maintain percentage constancy over a fairly wide range of background numbers. The constancy of some of the items at the top and bottom of the list, words such as 'all,' 'everybody,' 'practically everybody,' and 'none' and 'nobody,' might seem at first
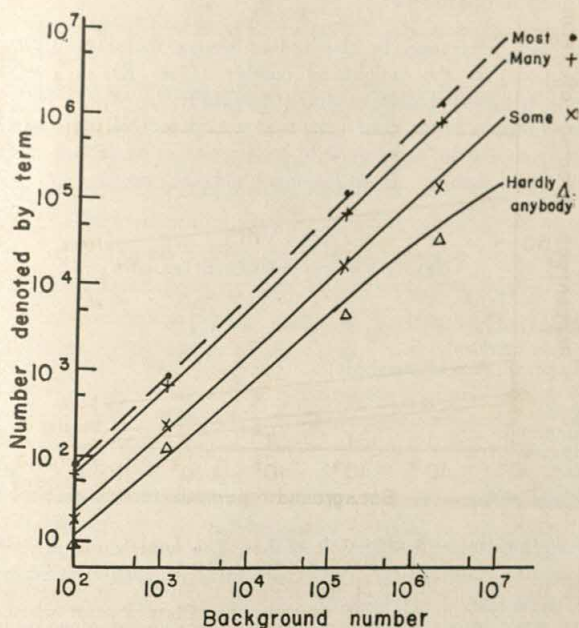
FIG. 3. NUMBERS DENOTED BY CLASS TERMS

The keying symbols used to plot the observed meanings are indicated at the right. The term 'most' is included as an example of a Class R term whose absolute meaning is proportional to the background number. The curves are computed on the basis of Equations [2] and [3].

sight to contradict the hypothesis. Closer consideration shows that it does not since, according to AL theory, stimuli farthest removed from AL tend to be less affected by changes in AL than are stimuli in its vicinity. Since percentage grand means of the estimates, which may be taken to represent the AL, were between 39.0 and 43.5%, stimuli near the two ends of the list would not be expected to show much variation with background number. The greatest change in quantitative denotation of the words was found with the intermediate terms in the list and this is entirely in line with the presuppositions underlying the study.

While we do not assert that this study alone establishes the validity of this technique for the study of the meaning of terms or furnishes a com-

plete justification of our interpretation, it does suggest that further work along these lines may yield results of more than passing interest. To the extent that the groups of Ss employed here are typical of the general population, certain facts have been established regarding the usage of non-numerical quantitative terms:

(1) Many such terms may be classified as having meanings which are either purely proportional to the background number (Class R) or intermediate between proportionality and absolute number (Class AR).

(2) Class A words, while their percentage meanings shift with number, retain a constant difference from background to background in percentage meaning.

(3) Many terms, such as 'all,' 'a majority,' 'nobody,' are used in senses appreci-

TABLE VIII

TERM STRENGTHS OF CLASS R TERMS

Term	$\log_e X$	Strength (X)
1, 2. All (Everybody)	1.43	4.18
3, 4, 5. Practically everybody (Almost all, Almost everybody)	1.22	3.39
6. Most	1.02	2.78
7. Generally	0.94	2.56
8. A majority	0.70	2.03
13. Fairly common	0.54	1.72
16. Too few	−0.14	0.87
25, 26. Nobody (None)	−0.48	0.62

ably different from their dictionary definitions. The term 'a majority' is especially interesting in this respect, since the average value of 60% denoted by this term carries interesting political implications.

(4) Even terms ordinarily considered to be vague in meaning, such as 'a considerable number' or 'some' have reasonably precisely definable meanings.

(5) A technique is available for showing the existence of synonyms, or for demonstrating that terms frequently classed as synonyms are not used synonymously. Thus 'many' represents about the same number in 150,000 as does 'a considerable number,' but the difference in behavior of the two terms prevents us from calling them synonyms.

Further tests of terms of the type employed here, with different populations, are still needed. It would also be interesting to develop similar techniques for classifying and studying the usage of terms of other types, for example, adjectives used to express condemnation or commendation. Studies of this sort should be of both semantic and social interest. Individual departures from normal usage may have diagnostic value.

PRACTICE AND TRANSFER IN THE VISUAL AND AUDITORY RECOGNITION OF VERBAL STIMULI

By Leo Postman and Mark R. Rosenzweig, University of California

The recognition of verbal stimuli constitutes a strategic area of contact between the fields of perception and verbal learning. Frequency of exercise, which is the basic variable in verbal learning, has been shown to be a major correlate of visual recognition-thresholds for words[1] and a similar relationship has been reported for the auditory discrimination of spoken messages.[2] The factor of recency has been shown to be a significant determinant of visual recognition,[3] and conditions of training conducive to intraserial interference have been found to have an adverse effect on the visual recognition of members of the series.[4]

Prior usage encompasses a complex of variables whose effects on recognition-thresholds cannot be disentangled as long as sheer frequency is used as the independent variable. First, frequency of exercise may refer either to number of exposures to a stimulus or to number of responses to a stimulus. There is some experimental evidence indicating that it is frequency of response rather than sheer frequency of exposure which is significantly related to recognition-thresholds.[5] Secondly, the question arises as to how specific the effects of exercise are to the sense-modality used during the training. Previous methods of measuring and manipulating frequency have not permitted the separation of the effects of training in different sense-modalities upon recognitive thresholds. The Thorndike-Lorge word-counts represent the frequencies with which words occur in *written* English, and they may be used to estimate the effects of prior reading on visual recognition-

[1] See, for example, D. H. Howes and R. L. Solomon, Visual duration threshold as a function of word probability, *J. exp. Psychol.,* 41, 1951, 401-410; Leo Postman and B. M. Schneider, Personal values, visual recognition and recall, *Psychol. Rev.,* 58, 1951, 271-284. For the word-counts used in these experiments see E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words,* 1944. In other studies, frequency of exercise has been controlled experimentally, *e.g.* Solomon and Postman, Frequency of usage as a determinant of recognition thresholds for words, *J. exp. Psychol.,* 43, 1952, 195-201; Patricia King-Ellison and J. J. Jenkins, The durational threshold of visual recognition as a function of word-frequency, this JOURNAL, 67, 1954, 700-703.
[2] Howes, The intelligibility of spoken messages, this JOURNAL, 65, 1952, 460-465.
[3] Postman and Solomon, Perceptual sensitivity to completed and incompleted tasks, *J. Personal.,* 18, 1950, 347-357.
[4] Postman, Experimental analysis of motivational factors in perception, in the Nebraska Symposium, *Current Theory and Research in Motivation,* 1954, 78 f.
[5] Leo Postman and Beverly Conger, Verbal habits and the visual recognition of words, *Science,* 119, 1954, 671-673.

thresholds. It must be noted, however, that words also have varying frequencies of occurrence as auditory stimuli. To the extent that (*a*) the frequencies of reading and hearing are not perfectly correlated and (*b*) the frequency of hearing has significant effects on visual recognition, the correlation between word-counts and visual recognition-thresholds is reduced. When frequency of usage has been manipulated experimentally, both visual and auditory training have been used simultaneously.[6] The question remains open as to the relative weight of the two types of training in producing variations in visual thresholds.

The present study attempts to refine and extend the analysis of the variable of exercise as a determinant of the thresholds of recognition for verbal stimuli. Specifically, two questions will be considered: (*a*) Does frequency of past exercise exert comparable effects on the auditory as well as the visual recognition of verbal stimuli? (*b*) To what extent are the effects of exercise specific to the sense-modality exercised, *i.e.* what is the degree of transfer from visual training to auditory recognition and conversely? Answers to these questions should help us to identify the mechanism of the effect of exercise on recognition.

METHOD

Four experimental groups and two control groups were studied. The experimental groups were given controlled experience with sets of nonsense-syllables and then tested for recognition of these and other syllables. In each group, recognitive thresholds could be related to frequencies of prior exercise. There were two conditions of training—visual and auditory—and two conditions of testing—again visual and auditory. A different group of *S*s was used for each of the four possible combinations of training and testing. The four experimental conditions will be referred to as *V-V* (visual training, visual test), *A-V* (auditory training, visual test), *A-A* (auditory training, auditory test), and *V-A* (visual training, auditory test). The *V-V* and *A-A* conditions will be designated as 'single-modality' conditions; *A-V* and *V-A*, as 'dual-modality' conditions. The two control groups had no preliminary training; one was given the visual test and the other was given the auditory test.

Selection of critical stimuli. For maximal efficiency in measuring the effects of practice and transfer, the syllables were selected to meet a number of requirements. The different sets of syllables were made as comparable as possible with respect to (*a*) frequency of occurrence as syllables in written English; (*b*) frequency of occurrence as three-letter combinations ('trigrams') in written English; (*c*) visual recognition-thresholds prior to training; and (*d*) auditory recognition-thresholds prior to training. Furthermore, to facilitate the measurement of transfer from one modality to another, the syllables were so selected that (*e*) a high percentage of *S*s spelled the items correctly when the syllables were dictated, and (*f*) a high percentage of *S*s chose the correct pronunciation when the syllables were presented visually.

[6] H. Ohms, Untersuchung unterwertiger Assoziationen mittels des Worterkennungsvorganges, *Z. Psychol.,* 56, 1910, 1-84; Solomon and Postman, *op. cit.,* 195-201.

Frequency of occurrence of syllables in written English was obtained from Dewey's count; the frequency-counts were converted into percentiles and the syllables were then dichotomized as of high or low frequency of occurrence.[7] Similarly, we dichotomized the trigram-frequencies obtained from Pratt's count into a high and a low group.[8] Of course, only syllables appearing in both Dewey's and Pratt's counts could be used. Visual and auditory recognitive thresholds were then determined for a large sample of these syllables. Visual thresholds were determined here, as in the main experiment, by varying the intensity of illumination in tachistoscopic presentation. Auditory thresholds were measured, again as in the main experiment, by presenting the syllables with various intensities of masking noise. To obtain the necessary information concerning spelling, the syllables were dictated to some groups of Ss who were instructed to spell them. To obtain information concerning pronunciation, the syllables were shown to other groups of Ss who were instructed to choose between alternative pronunciations presented on a tape recording. On the basis of these procedures, a final list of 10 *critical syllables* was chosen for use in the main experiment. This list consists of five pairs of syllables, one member of each pair being high on both the syllable- and trigram-counts, and the other member being low on both counts. The critical syllables did not vary widely in visual or in auditory discriminability. All the syllables satisfied rigorous criteria with regard to S's ability to spell and pronounce them correctly.

The five pairs of critical syllables are the following (the high frequency member of each pair being listed first in each case): (1) *DIS, SUG;* (2) *ING, DEV;* (3) *EST, VIT;* (4) *VAL, TUS;* (5) *VES, FAM.*

Training materials. To give differential exercise to the critical syllables, S was presented with a training series comprising the critical syllables embedded in a list of other nonsense-syllables taken from Dewey's tables. The frequency with which the different critical syllables appeared in the training series was varied systematically. The members of different pairs were presented with frequencies of 15, 5, 2, 1 and 0, respectively. Thus, there were 46 critical items scattered through the training series; in addition, there were 26 fillers which occupied fixed positions in the training series. Each of the five pairs of syllables was used equally often at each frequency of exercise, necessitating five different forms of the training series. In each form, the items for a particular frequency of exercise occupied the same positions.

Test-materials. The test-series consisted of 20 items: the 10 critical nonsense-syllables, 8 English three-letter words, and 2 additional nonsense-syllables (*PRE* and *GAR*).[9] The order in which the critical syllables and English words were presented was randomized for successive trials of the test.

The English words and their frequencies of usage according to the Thorndike-Lorge G (general) count were as follows: *BUN (4), FIN (6), HEM (14), LAP (36), MAD (A), PEN (A), PUB (1), RIB (31).*

---

[7] George Dewey, *Relativ [sic] Frequency of English Speech Sounds,* 1923, 60-82.
[8] Fletcher Pratt, *Secret and Urgent,* 1939, 264-277.
[9] Either an English word or one of these two additional syllables always appeared in the first and last positions of the test series. Since first position might be disadvantageous because of a 'warm-up' effect, and since responses made to the last item might be carried over to the next series, it seemed desirable to exclude critical items from the first and last positions.

Conditions of training. The *S*s receiving visual training were informed that *E* was concerned with the effectiveness of repetition in learning to write 'unfamiliar' words correctly. The stimulus-words, printed in capital letters, were presented on 2 × 2 in. slides. An *SVE* projector, with a 150-w. lamp, was used. Each item was projected on a milk-glass screen for 1 sec. This duration of exposure was fully adequate for correct recognition of the items. As soon as a word had been presented, *S* printed it in the appropriately numbered blank of a record-sheet, using a strip of cardboard to cover all spaces of the record-sheet except the one in which he was writing.

In the case of auditory training, the procedure was presented as an experiment on the effectiveness of repetition in learning to pronounce 'unfamiliar' words correctly. The items of the training series were presented by means of a tape-recording. The presentation of each word was followed by the signal, 'Now,' upon which *S* repeated the word aloud, under instructions to pronounce the word exactly as the speaker had. *S* was informed that his pronunciation was being monitored.

The order of presentation was the same for the visual and auditory training. All groups were informed that some words would occur more often than others, but no instructions to learn any of the words were given. The *S*s were run in small groups, ranging from three to six in number.

Recognition-tests. The recognition-test followed immediately after the training series. The test was introduced as a second and different experiment in the study of linguistic behavior, in which *E* was concerned with the relative ease of perception of 'English and foreign words' under difficult conditions of discrimination. No mention was made of the fact that some of the items of the training series would be used in the test.

The visual test consisted of tachistoscopic presentation of the stimulus-items. A photographic shutter was attached to the lens of the projector and set for exposures of 0.01 sec. throughout the test. Recognition-thresholds were determined by means of variation in the intensity of the stimulus-flash. The list of test-items was presented seven times, each time in a different random order. The intensity of the stimulus-flash was increased on successive test trials. A measure of contrast will be used to specify the stimulus-conditions for successive trials. Percentage of contrast is defined as $100(L_1 - L_2)/L_1$, where L_1 is the luminance of screen, and L_2 is the luminance of stimulus-letters. Intensity of illumination was changed by means of a variac, and luminance was measured in foot-lamberts with a General Electric Brightness Meter. The percentages of contrast for the successive presentations of the test-series were as follows: (1) 9.0, (2) 13.8, (3) 15.7, (4) 18.5, (5) 20.6, (6) 24.4, (7) 27.0.

The auditory test involved presentation of the test-items with different degrees of masking noise. The test-items and the noise, mixed electronically, were recorded on tape. The list of test-items was presented six times, each time in a different random order. The intensity of the speech was held constant throughout. The intensity of the masking sound (white noise) was reduced for successive presentations of the series. The signal-to-noise ratio will be used to specify the stimulus-conditions for successive trials. Signal-to-noise ratio is defined as the intensity of the speech-sound — the intensity of the masking noise. These intensity-levels were measured in db. with a General Radio Sound Level Meter. The signal-to-noise ratios (in db.) for the successive trials were as follows: (1) —7, (2) —3, (3) +1, (4) +7, (5) +13, (6) +19.

For both the visual and auditory tests, S was provided with record-sheets numbered from 1 through 20. Separate sheets were used for each presentation of the series. A test-item was always preceded by announcement of its number, and S used a cardboard strip to cover all the spaces of the record-sheet except the one in which he was to write. The experimental sessions were held in a light-proof and sound-shielded room.

Subjects. The Ss were 150 undergraduate students. There were 25 Ss in each of the four experimental groups and in each of the two control groups. The Ss did not know the purpose of the experiment and they were assigned to the experimental conditions at random. At the end of the experiment, S was asked to note his native language on the test-blank. Those Ss whose native tongue was not English were discarded from the sample and replaced. An S was retained only if his score on the last trial of the test was 80% or better. Actually, those rejected scored less than 50% correct on the final test-series, and it was felt that their records were atypical and would not provide reliable measures of discrimination. Since this criterion resulted in the rejection of a larger number of Ss from the untrained control groups than from the experimental groups, the results of the experimental and control groups are not comparable with respect to the absolute level of performance.

RESULTS

(1) Recognition-thresholds. Our measure of threshold will be the number of presentations required for the first correct recognition of an item. Although the successive test-trials do not represent equal changes in stimulus conditions, this measure provides us with a convenient rank-ordering of the items with respect to ease of recognition. Another measure, the number of correct recognitions of an item throughout the test-series, gave strictly equivalent results which will, therefore, not be reported.

(a) Thresholds for critical syllables. Fig. 1 shows the average recognition-thresholds for the critical nonsense-syllables as a function of frequency of exercise for the four experimental groups. In all cases, the thresholds fall with increasing frequency of exercise. The amount of variation produced by prior exercise and the shape of the function differ from condition to condition. When the test is visual, frequency of prior visual training $(V\text{-}V)$ significantly influences the recognition-thresholds $(F = 3.47, df. = 4$ and $180, P < 0.01)$, but the thresholds drop rather slowly with exercise, and the only significant break occurs between the frequencies of 5 and 15.[10] The visual test following the auditory training $(A\text{-}V)$ shows only small and insignificant differences in threshold as a function of frequency $(F = 0.82)$. The single-modality group $(V\text{-}V)$ has lower thresholds than the dual-modality group $(A\text{-}V)$ at all frequencies of exercise. The two groups differ significantly with respect to the level of their thresholds of recognition $(F = 10.13, df. = 1$ and $360, P < 0.01)$. Although the variable of frequency was significant for $V\text{-}V$ and not for $A\text{-}V$, the slopes of the two curves are not significantly differ-

---

[10] To determine critical breaks between conditions, we followed the procedure described by J. W. Tukey (Comparing individual means in the analysis of variance, *Biometrics*, 5, 1949, 99-114).

ent from each other. The interaction term, Condition × Frequency, yields an F of 0.81.

With the auditory test, both conditions of training produce significant variations in threshold. For A-A, $F = 10.59$ ($df. = 4$ and 180, $P < 0.01$); significant breaks occur between the frequencies of 0 and 1, and between 5 and 15. For the V-A
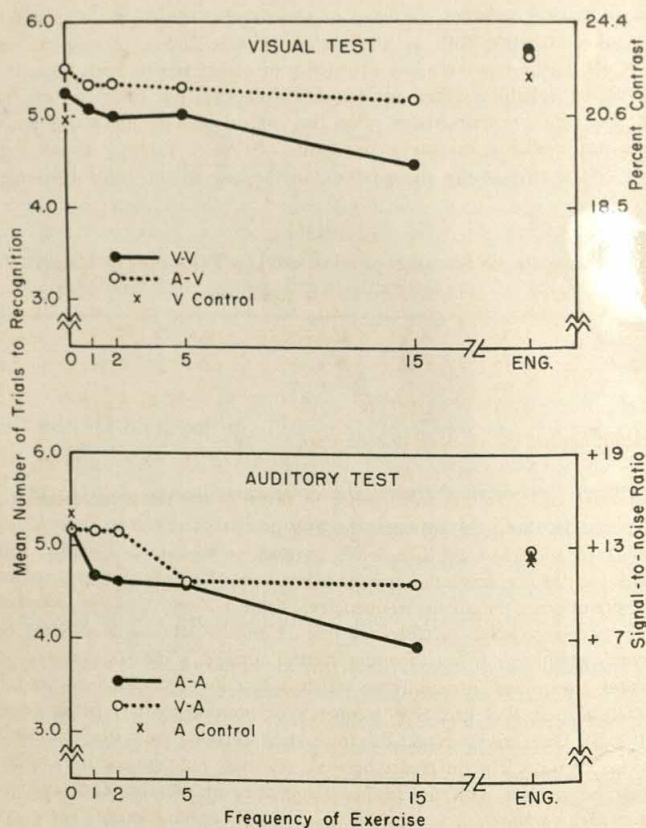
FIG. 1. MEAN NUMBER OF TRIALS TO RECOGNITION OF CRITICAL SYLLABLES AND ENGLISH WORDS FOR THE VARIOUS GROUPS
(Physical conditions of stimulation corresponding to test-trials are indicated on the right ordinate.)

condition, $F = 6.13$ ($P < 0.01$); a significant break occurs between the frequencies of 2 and 5. As before, the single-modality group shows lower thresholds than the dual-modality group ($F = 17.61$, $df. = 1$ and 360, $P < 0.01$). In addition, the slopes of A-A and V-A differ significantly ($F = 2.74$, $df. = 4$ and 360, $0.01 < P < 0.05$).

The determinants of the speed of recognition exert their effects throughout the successive presentations of the test-series. When psychometric functions are plotted for exercised syllables, unexercised syllables, and English words, typical sigmoid curves are obtained, with the levels of the functions consistently reflecting the average values of the thresholds.

(b) Linguistic frequency as a determinant of threshold. As we have mentioned, each pair of critical syllables consisted of one member which has a relatively high frequency of occurrence both as an English syllable and as a trigram, and one member which has a relatively low frequency of occurrence in both respects. These two groups of syllables differ in frequency of exercise prior to the training given to the experimental groups. For the control groups these differences are, of course, not modified by any experimental training. Table I shows the mean thresholds of the two groups of syllables. In the case of the control groups, there

TABLE I

Thresholds for Syllables of High and Low Frequency of Usage

Frequency	Condition					
	V-V	A-V	V-control	A-A	V-A	A-control
High	4.57	5.06	4.66	4.24	4.75	5.09
Low	5.22	5.53	5.30	4.65	5.03	5.62
t	3.94*	3.55*	3.81*	3.22*	3.11*	6.54*

* Significant beyond the 1% level of confidence.

is a substantial and significant difference in favor of the relatively familiar items. The difference in threshold between the two groups of syllables appeared despite our attempt to pre-select syllables with comparable thresholds. Is this difference maintained for the experimental groups when controlled practice is superimposed upon the pre-existing linguistic frequencies? Table I shows that the difference between the two groups of syllables is just as substantial and significant for the experimental groups as it was for the control groups. (The comparisons for the experimental groups are based only on syllables that were presented during training, *i.e.* the critical pair that had zero frequency of exercise for any given group was not included.) One might expect an interaction between linguistic frequency and experimental exercise. The difference between relatively familiar and unfamiliar items might well be smaller with the higher frequencies of exercise. Our experimental design does not provide for a sensitive test of such an interaction. For a given *S,* there was only one familiar and one unfamiliar syllable at each frequency of exercise. Perhaps for this reason no clear or significant evidence for interaction appears in our data. It is clear, however, that for all frequencies combined, the effects of linguistic usage remain intact.

Quite independently of linguistic frequency, some syllables benefit more from experimental exercise than do other syllables. Significant interactions between syllables and experimental frequency are found under all the experimental conditions. We have been unable to detect any systematic properties of the syllables which would help explain these interactions and are, for the time being, forced to ascribe them to structural (visual or phonetic) properties of the stimuli.

The analysis of the threshold-data supports the following conclusions: (*a*) frequency of exercise is a significant determinant of threshold; (*b*) larger effects of training are found with the auditory test than with the visual test; (*c*) larger effects of training are found for the single-modality conditions than for the dual-modality conditions; (*d*) following experimental exercise, syllables of high linguistic frequency continue to have lower thresholds than syllables of low frequency.

(c) Differences in sensitivity, The differences between visual and auditory tests may be due in part to certain features of the two methods of testing: (*a*) Examination of the psychometric functions for individual *S*s shows that the curves often are steeper under visual than under auditory conditions. In vision, recognition frequently was an all-or-none matter, and the determination of the threshold was thereby rendered difficult. The *S*s varied widely in their tachistoscopic thresholds. The steepness of the psychometric functions may have resulted from the fact that the changes in stimulus-intensity were chosen to accommodate groups of *S*s and were not optimal for measuring the thresholds of all individuals *S*s. The values of the masking noise used in the auditory test did not appear to suffer from such a limitation to a comparable degree. (*b*) Auditory discrimination in our test room was relatively unaffected by the position of *S* with regard to the speaker. Visual discrimination, on the contrary, was markedly affected by the position of *S* with regard to the screen. This fact tends to make the effects of the experimental variables stand out more clearly in the auditory than in the visual test. (*c*) On the other hand, performance on the auditory test may have been somewhat handicapped by the fact that *S* had to reproduce by visual symbols what he had perceived auditorily. Furthermore, the symbols have the disadvantage of being alphabetic rather than phonetic. To some extent, this handicap was overcome by the selection of syllables in terms of *S*'s ability to spell them correctly.

(d) Thresholds for English words. The mean thresholds for English words are indicated at the right-hand side of Fig. 1. The English words were not recognized more readily than the nonsense-syllables. In fact, the groups tested visually had higher thresholds for the English words than for the nonsense-syllables that received no exercise at all. In the auditory test, the English words behaved like nonsense-syllables of low frequency of exercise. Considered as syllables, the English words on the average have lower linguistic frequencies than do the nonsense-items, but the fact that they are complete and meaningful units of speech might make *S* more ready to use them as responses. The failure of the English words to yield lower thresholds than the nonsense-syllables suggests that *S* is no less ready to use syllables as response-units than he is English words of comparable linguistic frequency.

In the case of English words, the thresholds and psychometric functions for the single-modality and dual-modality groups are quite similar; that is, the difference found with nonsense-syllables is not paralleled with the English words. Thus, the effects of the preliminary training are specific to the material exercised, *i.e.* critical nonsense-syllables.

(2) Completion of partially discriminated items. One way in which prior exercise can serve to lower the recognition-threshold is by reducing the degree of discrimination required for correct reproduction of an item. Under reduced conditions of stimulation, such as tachistoscopic

presentation or auditory presentation with masking noise, only fragments of the stimulus may be correctly discriminated on any given occasion, *e.g.* only one or two letters of a nonsense-syllable. The more familiar the item, the smaller the fragment of the total stimulus needed for correct identification of the item. Another way of stating the same point is to say that the preliminary training reduces the number of alternatives in terms

TABLE II

INDEX OF TENDENCY TO COMPLETE ITEMS

(Entries are mean differences between the numbers of complete responses containing three correct letters and two correct letters respectively)

Frequency of exercise	Condition			
	V-V	A-V	A-A	V-A
15	13.8	12.4	11.7	7.8
5	10.7	10.3	6.5	7.6
2	11.5	11.3	7.8	3.9
1	11.2	10.6	7.6	4.6
0	9.7	8.9	4.7	4.8
English words	5.6	6.6	4.6	4.5

of which *S* can interpret fragments which he has recognized. These considerations suggest an additional measure which may be used for evaluation of the effects of prior exercise. This measure is derived from an examination of *S*'s complete responses, *i.e.* responses containing three letters. Let us consider (a) those complete responses which contain two correct letters out of three, and (b) those complete responses in which all three letters are correct. (Except for rare cases of transposition of the letters, the latter cases will be correct recognitions.) The higher the frequency of prior exercise, the larger should be the number of complete responses that are correct relative to the number of complete responses containing only two correct letters. If the item is relatively familiar, correct discrimination of two letters should lead readily to the identification of the item, resulting in a correct response. If, on the other hand, the item is relatively unfamiliar, discrimination of two of the letters is less likely to lead to a correct response. The average difference between the number of complete responses containing three correct letters and the number of complete responses containing two correct letters was determined for the critical nonsense-syllables and English words. These measures may be regarded as an index of the reduction of alternatives by prior training or as an index of the tendency toward correct completion. The results for the experimental groups are presented in Table II.

For all conditions, the tendency toward completion increases as a

function of frequency of exercise. The patterns of variation in the index closely parallel those observed with thresholds of recognition. The tendency to complete items in terms of a reduced number of alternatives constitutes at least one of the mechanisms by which preliminary exercise leads to a lowering of recognition-thresholds. For each of the conditions, an analysis of variance was performed to determine the significance of changes in the index of completion. The changes in the index are significant for $A\text{-}A$ ($F = 6.47$, $df. = 4$ and 96, $P < 0.01$) and for $V\text{-}A$ ($F = 2.52$, $P = 0.05$). The F-ratios are not significant for either of the groups tested visually.[11] It will be recalled that there was no significant variation in recognition-thresholds with exercise for $A\text{-}V$ but that thresholds did vary significantly for $V\text{-}V$. It appears that the index of completion is less sensitive to the variable of frequency than is the threshold of recognition. This finding is reasonable if we assume that tendency toward completion is only one of the ways in which preliminary training exerts its effects.

As in the case of recognition-thresholds, we find that the effects of training are considerably more clear-cut when the test of recognition is auditory than when it is visual. One of the major reasons for this difference between the sense-modalities appears to lie in the fact that auditory presentation results in a considerably higher proportion of complete three-letter responses than does visual presentation. Evidence for the greater tendency toward complete three-letter responses to auditory stimuli is presented in Fig. 2. This figure shows, for the different conditions, the percentages of all incorrect responses that consisted of one, two, and three letters respectively. The percentages are based on the total frequencies of error, excluding failures to respond.[12] It is clear that auditory presentation results in strikingly higher percentages of three-letter responses than does visual presentation. Auditory discrimination of verbal stimuli tends to be made in units such as syllables. Visual discrimination is more analytic, and S can effectively attend to units of different sizes.

(3) Development of pre-recognition responses. The role of S's verbal habits in determining thresholds of recognition is further clarified by an examination of the frequencies of different types of error at various stages prior to correct recognition. We shall limit this examination to erroneous

---

[11] Although the F-ratios were not significant for the two visual test-conditions, the differences between the two extreme points of each of these conditions—0 and 15 repetitions—were significant at the 1% level of confidence. While these tests indicate that exercise was effective here too, such selected t-tests must be interpreted cautiously.

[12] In analyzing the errors, we have omitted from consideration responses to the 'buffer' syllables, *PRE* and *GAR*.

responses consisting of three letters, *i.e.* responses which are complete but incorrect. Such errors may be regarded as wrong guesses or 'hypotheses,' often based on the discrimination of stimulus-fragments.

(*a*) *Hypotheses.* We have divided the pre-recognition hypotheses into English words and nonsense-syllables. The total frequencies of the two types of 'hypotheses'
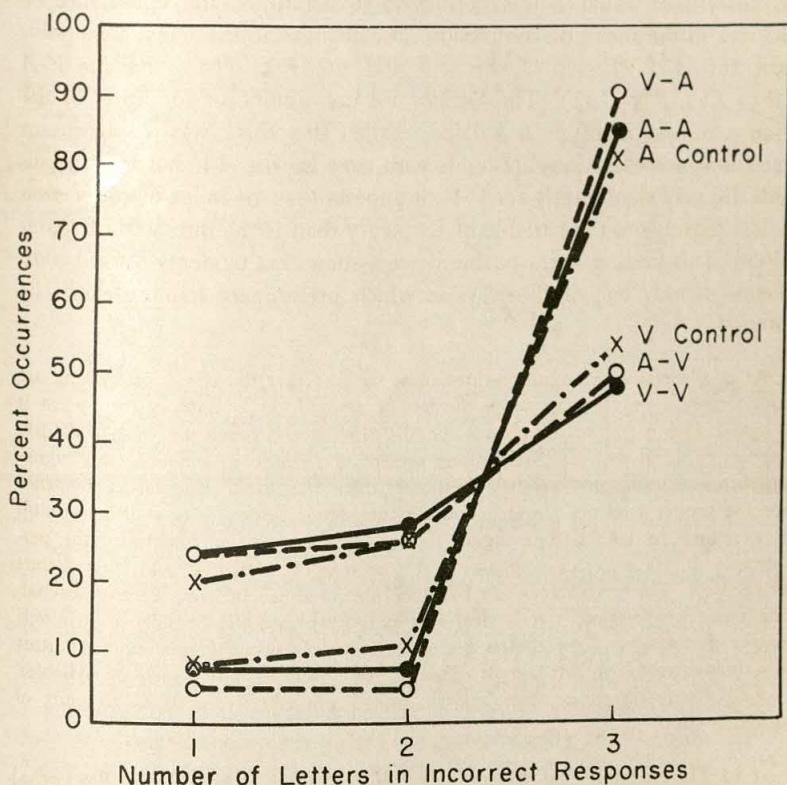
FIG. 2. PERCENTAGES OF INCORRECT RESPONSES CONSISTING OF ONE, TWO, AND THREE LETTERS

for English and nonsense syllables given by the different groups are presented in Table III. The over-all frequencies of hypotheses (*N* and *E* in Table III) are considerably higher for the auditory test than for the visual test. This finding is in agreement with the tendency, already noted, for more frequent complete responses in the auditory than in the visual test. Incomplete (one- or two-letter) responses and blanks (failures to respond) are relatively much more frequent for the visual test than for the auditory test. Under all conditions, English-word 'hypotheses' occur more frequently than nonsense-syllable 'hypotheses.' The *relative* frequency of

the English-word 'hypotheses' varies, however, from condition to condition. The percentages of nonsense-syllable 'hypotheses' are larger for the single-modality conditions (*V-V* and *A-A*) than for the dual-modality conditions (*A-V* and *V-A*) and control conditions. The higher frequency of nonsense-responses given by the single-modality groups may be attributed to the effects of preliminary training. Having been exposed to a series of nonsense-syllables under conditions identical with those of the test, an *S* in a single-modality group is more ready to respond with nonsense-syllables than is an *S* for whom the conditions of training

TABLE III

FREQUENCIES OF ENGLISH- AND NONSENSE-HYPOTHESES PRIOR TO
CORRECT RECOGNITION

Hypotheses	Condition					
	V-V	A-V	V-control	A-A	V-A	A-control
Nonsense	210	148	196	403	381	286
English	242	204	278	513	586	619
$N+E$	452	352	474	916	967	905
$100(N/N+E)$	46.5	42.0	41.4	44.0	39.4	31.6

and testing are different, or who has not had any preliminary training. When the test is auditory, the dual-modality group (*V-A*) has a higher percentage of nonsense-responses than the control group. When the test is visual, the percentages of nonsense-responses for the dual-modality group (*A-V*) and the control group are approximately the same. This finding confirms the conclusion that the visual test is less sensitive to the effects of preliminary training than is the auditory test. As far as the pre-recognition 'hypotheses' are concerned, the visual test does reflect preliminary visual training but is not sufficiently sensitive to reveal transfer from auditory training. The auditory test, on the other hand, reflects both types of preliminary training.

The frequency of pre-recognition 'hypotheses' is relatively low for the early and late test-trials and rises to a peak toward the middle of the test-series. In the earliest trials of the test, when the conditions of discrimination are very poor, there are few 'hypotheses,' since failures to respond account for the majority of *S*'s reactions at that stage. As the test-trials continue, and the discrimination of stimulus-fragments becomes possible, the frequency of incorrect complete responses increases steeply; they represent unsuccessful attempts at completion by *S*. As the test approaches its end, correct recognitions steadily increase in frequency and the opportunity for errors is correspondingly reduced. It is clear that the trials in the middle of the series, which may be regarded as the threshold-region, are the most sensitive to *S*'s sets and verbal habits. It is, indeed, these trials which reveal the most consistent differences among conditions with respect to the frequency of pre-recognition hypotheses.

(*b*) *Intrusion of exercised critical syllables.* The *S*'s disposition to respond with exercised items shows itself most directly in the intrusion of critical syllables as incorrect responses during the test. Since each critical syllable shared at least one letter (in the auditory case, one phoneme) with other items in the test-series, the stimulus-conditions were conducive to such intrusions. The control groups did

not give a single intrusion of a critical nonsense-syllable. Occurence of such intrusions in the experimental groups must, therefore, be ascribed to the effects of preliminary training; they cannot be due to structural characteristics of the stimuli or to whatever learning may take place during the test itself. The absolute number of intrusions was rather small and the distributions showed pronounced positive skewness.

The median numbers of intrusions for the experimental groups were as follows: V-V, 1.44, A-V, 0.62, A-A, 1.25, V-A, 0.96. For both tests, the single-modality group showed a larger median number of intrusions than did the dual-modality group. As was found to be true of hypotheses in general, the largest number of intrusions tended to occur toward the middle of the list. The peak-frequency of intrusions occurred earlier in the test-series for the single-modality groups than for the dual-modality groups. Both the greater median number and the earlier peak of intrusions in the single-modality groups suggest that intrusions are contingent on a lower degree of discrimination for the single-modality S than for the dual-modality S.

(4) Discrimination of word-classes. The discrimination of stimulus-elements appears to enable S to identify the items as being either English or nonsense even before he is able to respond correctly. This conclusion is supported by an analysis of the errors made in response to English- and nonsense-stimuli. Errors for the two types of stimuli were considered separately. For both English words and nonsense-stimuli, all errors were divided into three classes, those containing zero, one, and two correct elements (letters) respectively. For each of these classes of error, the numbers of English- and nonsense-hypotheses (incorrect three-letter responses) were then expressed as percentages of the total numbers of errors. Fig. 3 presents the differences between the percentages of English and nonsense 'hypotheses' given to English words and nonsense-syllables. These differences are plotted as a function of the number of correct elements. The data have been pooled for the three visual test-conditions and the three auditory test-conditions, since the results were similar for all groups receiving a given type of test.

The differences tend to be in favor of English 'hypotheses,' particularly for the auditory test-conditions (Table III). The relative predominance of English over nonsense 'hypotheses' is a joint function of (a) the number of elements correctly recognized, and (b) the nature of the stimulus. As the number of correct elements increases, the difference in favor of English hypotheses grows more rapidly when the stimulus is an English word than when the stimulus is a nonsense syllable. Examination of Fig. 3 shows that this differentiation of the responses to the two types of stimuli depends on the correct recognition of two elements; when fewer than two elements are correctly recognized, the relative frequencies of the two kinds of 'hypothesis' remain independent of the nature of the stimulus. It does appear as though S was able to distinguish between English words and nonsense-

syllables while he was still making errors about one of the letters. What mechanism can be invoked to account for this discrimination of the class of words? If we discard the possibility of 'subception,' which has no explanatory value, we might at first be tempted to conclude that S was discriminating between English words and nonsense-syllables in terms of the probabilities of letter-sequences in English. Some combinations of letters are more likely to occur in English than others. Having discriminated two of the letters, S might 'bet' on whether this
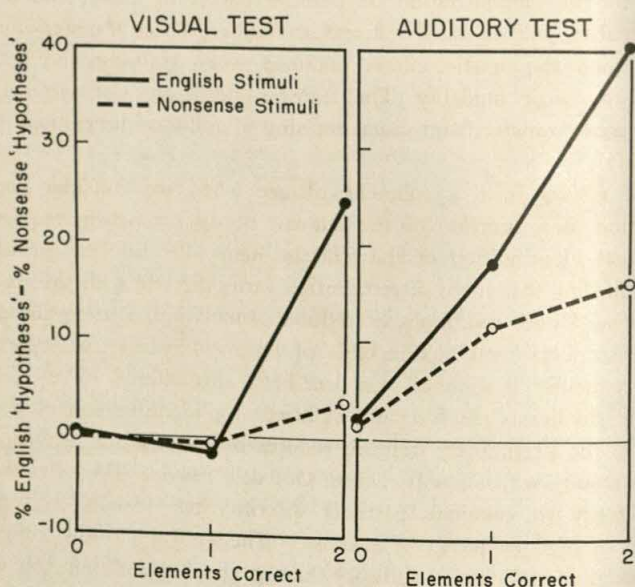
Fig. 3. Differences Between Percentages of Nonsense- and English-'Hypotheses' Given to Nonsense Syllables and English Words as a Function of Number of Elements Correctly Discriminated

combination was part of an English word. The frequency of usage of our critical nonsense-stimuli as *syllables* in English is, however, no lower than the frequency of the English words. Hence there is no difference between critical syllables and the English words in their conformity to the statistical structure of English. The same is true for two-letter sequences taken from the two types of stimuli. The answer to our problem turns out to be surprisingly simple. The structure of our critical syllables is such that once two of the three letters have been discriminated in their correct positions, it is often impossible to form a *three-letter English* word, and S must therefore, respond with a nonsense 'hypothesis' if he is to give a complete response. In the case of the English words, on the other hand, virtually any such pair of letters can be completed to form an English word other than the stimulus-item itself. Three-letter English words tend to share letter-combinations

to a greater degree than can be predicted from the statistical structure of English in general.[13]

DISCUSSION

In conformity with earlier studies, our experiment has shown that frequency of prior exercise is a significant determinant of the recognition-thresholds for verbal stimuli. The generality of this finding has been extended by the demonstration of parallel effects in visual and auditory discrimination and of transfer across modalities. These transfer-effects are smaller than the practice-effects obtained when training and test both involve the same modality. The transfer-effects are not symmetrical—there is more transfer from visual training to auditory discrimination than conversely.

These results form a coherent picture when we consider speed of recognition as a function of the amount of discrimination required for the correct identification of the stimulus-items. Let us first consider the general finding that speed of recognition varies directly with the frequency of exercise. Under conditions of reduced stimulation, S attempts to identify the stimulus-items on the basis of fragmentary discriminations. The more frequently a stimulus-item has been encountered or used in the past, the smaller is the fragment sufficient for identification of the total item, *i.e.* the preliminary training reduces the number of alternative responses among which S will choose. Our data provide direct evidence for S's tendency to complete partially discriminated stimuli in terms of syllables of high frequency of exercise.[14] The tendency toward completion was greater in response to auditory than to visual stimulation. Hence, the auditory test was more sensitive than the visual test to the reduction of alternative responses resulting from the preliminary training.

The reduction in alternative responses produced by the preliminary training does not, however, account for all of the variance in the thresholds. The relative frequency of usage in the language continued to exert significant effects. The syllables of relatively high frequency of usage were

---

[13] This conclusion was borne out by examination of samples of 25 additional three-letter English words and 25 additional three-letter English syllables.

[14] Several experimental studies have shown that familiarization with a list of English words also results in a considerable reduction in the recognition-thresholds. In such cases, the training served to reduce the number of alternative English responses among which S was prepared to choose. See, for example, J. C. R. Licklider and Irwin Pollack, Effects of differentiation, integration and infinite peak clipping upon the intelligibility of speech. *J. acoust. Soc. Amer.*, 20, 1948, 42-51; J. S. Bruner, G. A. Miller, and Claire Zimmerman, Discriminative skill and discriminative matching in perceptual recognition, *J. exp. Psychol.*, 49, 1955, 187-192.

recognized more readily than syllables of relatively low frequency of usage by all of the experimental and control groups. Experimental exercise did not eliminate the effects of prior familiarity. It may be recalled that syllable-frequency in English and trigram-frequency were tied together in classifying the stimuli. Since a previous experiment had failed to find a significant correlation between trigram-frequency and recognition-thresholds, it seems likely that frequency of occurrence as English syllables is responsible for the present effect.[15]

Next let us consider why the single-modality conditions show larger effects of training than do the dual-modality conditions. The readiness to identify an item on the basis of a partial discrimination depends, of course, on the perceived similarity between the recognized fragment and the total item. We would expect such identification to occur most readily when the same sense-modality is used in the training and the test. During the training, certain constellations of letters or sounds are presented repeatedly and become familiar to S. Clearly, a fragment discriminated by the same modality as that used during training will be more familiar than a fragment that is recognized via a different sense-modality.[16] When the modality remains the same, the responses made to an item during training are thus readily reinstated during the test. When the modalities are different, recognition of a familiar stimulus must depend on a mediating response which serves to 'translate' the input from one modality to the other. To the extent that the completion and recognition of items depends on the reinstatement at the time of test of reactions made during the training, it is plausible that the single-modality conditions show greater effects of preliminary training than the dual-modality conditions.

Analysis of the probable mediating mechanisms also helps to explain the fact that the transfer-effects are greater from vision to audition than in the opposite direction. A possible mediating mechanism is suggested by the fact that during visual training S tends to repeat stimulus-items subvocally. This subvocal repetition may mediate transfer to auditory discrimination. On the other hand, auditory training does not appear to produce reactions (e.g. visualizing of the items) which can similarly facilitate transfer across modalities. Our speculations here are in part confirmed by reports of the Ss. When questioned at the end of the experiment, Ss tended to state that they did pronounce to themselves items presented visually but did not visualize stimuli presented auditorily. This difference probably

---

[15] Postman and Conger, op. cit., 671-673.

[16] For a recent discussion of the effects of changes in modality upon recognition see Hans Wallach and Emanuel Averbach, On memory modalities, this JOURNAL, 68, 1955, 250.

does not exhaust the reasons for the asymmetry of transfer. As we have pointed out repeatedly, the auditory test *per se* is more sensitive to the effects of preliminary training than is the visual test and hence may be assumed to be more sensitive to transfer-effects. We conclude, nevertheless, that degree of transfer across modalities depends, at least in part, on communality of responses made to the stimulus-items during training and at the time of the test.

Another variable which may influence the effects of preliminary training and degree of transfer is the nature of the specific recording responses made by S during training and test. In our experiment, S receiving visual training wrote down each of the stimuli whereas S receiving auditory training repeated them aloud. During the test, all groups made their responses in writing. On the basis of the sheer formal similarity of the responses made during training and test, we should have to expect greater practice-effects for V-V than for A-A. Such is, of course, not the case; A-A shows consistently more clear-cut practice-effects than does V-V. If the identity of recording responses plays any part at all, this factor is clearly subordinate to the other conditions which make the auditory test more sensitive than the visual test to the effects of preliminary training. In the adult S, the equivalence of written and spoken responses is well established. The two types of response are roughly interchangeable, provided there is agreement on the spelling of items presented auditorily and on the pronunciation of items presented visually. Our items were selected to satisfy this condition. The systematic variations among the groups may, therefore, be ascribed primarily to the sensory conditions of training and testing.

Finally, our findings concerning transfer help in the interpretation of some earlier studies. In the experiment of Solomon and Postman, S looked at cards wih Turkish words and read the words aloud; the test was visual.[17] Our results suggest that in that experiment and in others using the same procedure, the visual component of the training was more important than the auditory component in changing the visual thresholds; however, the vocal-auditory exercise may have added to the effect. In studies of the recognition-thresholds of 'taboo' words, the Thorndike-Lorge count represents only frequencies of written usage whereas most of the usage of such words may be oral.[18] If there were no transfer of auditory training to visual recognition, the Thorndike-Lorge count would

[17] Solomon and Postman, *op. cit.*, 195-201.
[18] For a discussion of the problem of word-familiarity in relation to the recognition of taboo words see Howes and Solomon, A note on McGinnies' "Emotionality and perceptual defense," *Psychol. Rev.*, 57, 1950, 229-234.

be an adequate estimate of the frequency relevant to a visual recognition-test. Since we have found some small transfer-effects from hearing to vision, the frequency of written usage may leave out a relevant source of exercise. Similar considerations apply to all the studies in which exercise is measured in only one modality though it occurs in both.

SUMMARY

This experiment investigates the effects of preliminary training on the visual and auditory discrimination of verbal stimuli. A series of nonsense-syllables was selected for initial ease of visual and auditory discrimination. Different syllables were then given varying frequencies of exercise, ranging from 0 to 15 repetitions. Half the Ss received visual training; the other half were given auditory training. Following the training, recognition was tested. The test-series included the experimental syllables as well as English words. For half the Ss trained visually, the recognition-test was visual, while for the other half the test was auditory. Similarly, auditory training was followed either by a visual or by an auditory test. For visual recognition, the test-items were presented tachistoscopically under different intensities of illumination. For the auditory test, the items were presented in conjunction with different degrees of masking noise.

Frequency of past exercise was found to be a significant determinant of both visual and auditory recognition. The effects were more clear-cut in auditory than in visual discrimination. When the same sense-modality was involved in both training and test, the effects of practice were more pronounced than when there was a change in modality. The transfer-effects from visual training to auditory discrimination were more pronounced than conversely. The analysis of results stressed the reduction of alternative responses produced by the preliminary training. Auditory stimulation tended to produce more complete responses than did visual stimulation. The auditory test was more sensitive than the visual test to the reduction in the number of alternative responses resulting from the preliminary training.

ON THE SIMILARITY BETWEEN REACTIVE
INHIBITION AND NEURAL SATIATION

By Carl P. Duncan, Northwestern University

In 1950 Köhler and Fishback noted that certain empirical phenomena in the field of learning, such as the favorable influence of rest-intervals, and reminiscence, were very similar to empirical phenomena such as figural after-effects in perception.[1] They therefore suggested that processes hypothesized to account for these phenomena, viz. some kind of inhibition in learning and neural satiation in perception, were also similar. In the years since the publication of the Köhler and Fishback study more evidence pertinent to their suggestions has accumulated. The purpose of this paper is to review the relevant literature and to attempt to show that neural satiation and the particular kind of inhibition called reactive inhibition are highly similar in several respects. It is hoped that the comparisons to be drawn will help bridge the gap between the fields of learning and perception.

In 1943 Hull postulated that whenever a response occurs, there is developed a process called reactive inhibition, symbolized I_R.[2] Reactive inhibition was likened to a negative drive-state in that it reduces reaction-potential, and accordingly depresses performance. It was believed to accumulate in amount both as the number of responses, and as the amount of work involved in each response, increased. Hull formulated the reactive-inhibition postulate on the basis of a number of earlier observations and hypotheses, chief of which were the Mowrer-Miller hypothesis, Pavlov's several varieties of inhibition, and some of the studies on work, reminiscence, and distributed practice.[3]

Although reactive inhibition or a similar notion has been rather widely employed as an explanatory construct, it has been most intensively studied in situations employing human Ss performing on motor tasks. This literature is exemplified by studies of Ammons, Kimble, and many others on such motor learning tasks as the pursuit-rotor, block-turning, and reversed alphabet-printing, and by Bilodeau's studies with 'pure work' tasks such as cranking or dropping balls through a

chute.[4] With such tasks, reactive inhibition is postulated to account in part or entirely for the poorer performance under massed practice than under distributed practice, and the dissipation of reactive inhibition after rest is assumed to account for reminiscence. Most, if not all, of the relatively precise information about reactive inhibition which is needed for the arugment to be presented comes from this literature on motor tasks, so this paper will not review studies with verbal tasks or with animal *Ss*. Attention should also be called to the fact that reactive inhibition as a construct has recently been considerably improved by Ellis.[5]

The concept of neural satiation was elaborated most thoroughly by Köhler and Wallach.[6] Satiation was postulated to account for various figural after-effects which occurred both with some stimulus-figures studied earlier by Gibson, and with many other figures prepared by Köhler and Wallach themselves.[7] The typical procedure for observing figural after-effects in visual stimuli involves presenting an inspection-figure, which *O* fixates for a few minutes or less, followed by a test-figure presented in or near the place in the visual field formerly occupied by the inspection-figure. With appropriately chosen inspection- and test-figures, the test-figure will show distortions, as compared to some standard, in any one or all of the characteristics of color, size, and location. Köhler and Wallach believed that the basic distortion or after-effect was displacement of the test-figure from the region previously occupied by the inspection-figures.

Köhler and Emery have now found figural after-effects with three-dimensional visual figures,[8] Deutsch, Krauskopf, and others have obtained after-effects in audition, and Gibson, Köhler and Dinnerstein, Nachmias, and Jaffe, have demonstrated kinesthetic after-effects.[9] Figural after-effects are not, of course, limited to

---

[4] R. B. Ammons, Acquisition of motor skill: II. Rotary pursuit performance with continuous practice before and after a single rest, *J. exp. Psychol.*, 37, 1947, 393-411; G. A. Kimble, An experimental test of a two-factor theory of inhibition, *ibid.*, 39, 1949, 15-23; E. A. Bilodeau, Performance decrement in a simple motor task before and after a single rest, *ibid.*, 43, 1952, 381-390; Decrements and recovery from decrements in a simple work task with variation in force requirements at different stages of practice, *ibid.*, 44, 1952, 96-100.

[5] D. S. Ellis, Inhibition theory and the effort variable, *Psychol. Rev.*, 60, 1953, 383-392.

[6] Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

[7] J. J. Gibson, Adaptation, after-effect, and contrast in the perception of curved lines, *J. exp. Psychol.*, 16, 1933, 1-31; Adaptation, after-effect, and contrast in the perception of tilted lines. II. Simultaneous contrast and the areal restriction of the after-effect, *ibid.*, 20, 1937, 553-569; J. J. Gibson and Minnie Radner, Adaptation, after-effect, and contrast in the perception of tilted lines. I. Quantitative studies, *ibid.*, 20, 1937, 453-467.

[8] Wolfgang Köhler and D. A. Emery, Figural after-effects in the third dimension of visual space, this JOURNAL, 60, 1947, 159-201.

[9] J. A. Deutsch, A preliminary report on a new auditory after-effect, *Quart. J. exp. Psychol.*, 3, 1951, 43-46; John Krauskopf, Figural after-effects in auditory space, this JOURNAL, 67, 1954, 278-287; Gibson, *op. cit., J. exp. Psychol.*, 16, 1933, 1-31; Wolfgang Köhler and Dorothy Dinnerstein, Figural after-effects in kinesthesis, *Miscell. Psychol. Albert Michotte*, 1947, 196-220; Jack Nachmias, Figural after-effects in kinesthetic space, this JOURNAL, 66, 1953, 609-612; Robert Jaffe, Kinesthetic after-effects following cerebral lesions, this JOURNAL, 67, 1954, 668-676.

test-figures. Köhler and Wallach argue that satiation should also occur during fixation of original or inspection-figures, resulting in some distortion. They are thus able to give a plausible explanation for a number of illusions. Bales and Follansbee, Gibson, Weitz and Compton, and others, have also found changes in the inspection-figure during prolonged (5-10 min.) inspection-periods.[10] The presumed neurological basis of satiation has been criticized in some respects, especially by Smith, and Jaffe, and recently Osgood and Heyer have explained figural after-effects in terms of physiological mechanisms other than those postulated by Köhler and Wallach.[11] However, as in the case of reactive inhibition, apparently no one has denied the necessity for postulating some type of distorting process, so we shall continue to use the term satiation but without implying any particular neurological hypotheses.

SIMILARITIES BETWEEN REACTIVE INHIBITION AND NEURAL SATIATION

(1) Source. Both reactive inhibition and satiation seem to develop as a consequence of, but differ from, afferent stimulation. This appears obvious with regard to satiation, inferred as it is from behavior either during or following presentation of stimuli. It may not be so obvious with regard to reactive inhibition since the usual statement is merely that I_R develops whenever a response is made. It seems sensible to argue, however, that it is the kinesthetic stimulation arising from muscular contraction which largely causes the development of reactive inhibition (and perhaps some also develops from the visual stimulation too, in perceptual-motor tasks).

This suggestion is not completely original. In an extensive article Solomon argued, in effect, for the substitution of kinesthetic stimulation in place of reactive inhibition in Hull's postulate.[12] We are suggesting here that I_R is, like satiation, a central process—see Point *(2)* below—arising *from* afferent stimulation (kinesthetic in the case of motor tasks), not that it *is* stimulation. In this connection we may again call attention to the demonstrations of kinesthetic figural after-effects.[13]

(2) Locus. Köhler and Wallach, and Gibson, have shown that satiation has a central locus by demonstrating that figural after-effects occur when the inspection-figure is presented to one eye and the test-figure to the other eye.[14] Although Hull assumed that reactive inhibition was peripheral, *i.e.* confined to the effectors involved in responding, Ammons, Grice and Reynolds, Irion and Gustafson, Kimble,

---

[10] J. F. Bales and G. L. Follansbee, The after-effect of the perception of curved lines, *J. exp. Psychol.*, 18, 1935, 499-503; Gibson, *op. cit., ibid.*, 1933, 1-31; Joseph Weitz and Bertita Compton, A further stereoscopic study of figural after-effects, this JOURNAL, 63, 1950, 78-83

[11] K. R. Smith, The satiational theory of the figural after-effect, this JOURNAL, 61, 1948, 282-286; Jaffe, *op. cit., ibid.*, 668-676; C. E. Osgood and A. W. Heyer, A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118.

[12] R. L. Solomon, The influence of work on behavior, *Psychol. Bull.*, 45, 1948, 1-40.

[13] Gibson, *op. cit., J. exp. Psychol.*, 16,1 933, 1-31; Köhler and Dinnerstein, *op. cit.*, 196-220; Nachmias, *op. cit.*, 609-612; Jaffe, *op. cit.*, 668-676.

[14] Köhler and Wallach, *op. cit.*, 269-357; Gibson, *op. cit., J. exp. Psychol.*, 20, 1937, 553-569.

and Rockway, have shown that it is not confined to the particular effectors involved in the response.[15] Therefore, it may be assumed that I_R also has a central locus.

(3) *Effects.* Both processes have essentially the effect of distorting behavior away from some criterion or standard. This is obvious for satiation; the process is inferred from distortions in the perception of figures. In the case of reactive inhibition the effect while stimulation is still continuing, *e.g.* during highly massed practice on a motor task, is measured as a depression of performance as compared either to an initial performance-level or to performance of Ss working under distributed practice. (As noted earlier, the effect of reactive inhibition is also commonly measured in terms of recovery from distortion, *i.e.* reminiscence during rest following massed practiced.) The depression of performance must result, however, from a change along some dimension in which movement can vary.

In a speed task such as cranking, or dropping balls through a chute, the 'distortion' produced by accumulated I_R is a slowing of response-rate compared to the rate at the beginning of work.[16] In an accuracy-plus-speed task such as the pursuit-rotor, where performance is measured in terms of time on target, a depression of performance under massed practice is indicated by less time on target. This must mean that Ss' movements are distorted from the correct circular path. Since accuracy at a certain speed is required, the performance-depression may also be in part due to an incorrect speed.

There has been very little work done to study just what happens when performance is depressed under massed practice on a motor task requiring both speed and accuracy. There are, however, two studies which provide some information. Ammons compared the number of hits (contacts of stylus and target) in rotary pursuit performance under massed and distributed practice.[17] He found that the depression of performance (in terms of time on target) produced by massed practice was clearly reflected in the measures of hits. There were fewer hits and shorter mean duration of hits in the massed-practice group. Although this finding indicates that one of the effects of massed practice is to produce jerkiness of movements, it does not specify clearly enough the nature of the distortions.

Somewhat more useful findings were reported by Ammons, Ammons, and Morgan.[18] They took motion pictures of the movements made in rotary-pursuit performance under various conditions, one of which was massed versus distributed

---

[15] Hull, *op. cit.*, 281; Ammons, effects of pre-practice activities on rotary pursuit performance, *J. exp. Psychol.*, 41, 1951, 187-191; G. R. Grice and Bradley Reynolds, Effects of varying amounts of rest on conventional and bilateral transfer "reminiscence," *ibid.*, 44, 1952, 247-252; A. L. Irion and L. M. Gustafson, "Reminiscence" in bilateral transfer, *ibid.*, 43, 1952, 321-323; Kimble, Transfer of work inhibition in motor learning, *ibid.*, 43, 1952, 391-392; M. R. Rockway, Bilateral reminiscence in pursuit-rotor learning as a function of amount of first-hand practice and length of rest, *ibid.*, 46, 1953, 337-344.

[16] Bilodeau, *op. cit.*, *J. exp. Psychol.*, 44, 1952, 96-100; *ibid.*, 43, 1952, 381-390.

[17] Ammons, Effects of distribution of practice on rotary pursuit "hits," *J. exp. Psychol.*, 41, 1951, 17-22.

[18] R. B. Ammons, C. H. Ammons, and R. L. Morgan, Movement analysis of the performance of a simple perceptual-motor task under various conditions, *Technical Report 54-36, Wright Air Development Center*, 1954.

practice, and classified the movements into several types. The writer combined several of their classes of off-target movements into movements where the stylus-tip was either following or leading the target (by no more than their criterion of 60°), or was outside or inside the target (greater or lesser radius than the target). During the last practice-trial there were more following than leading movements, and more inside than outside movements, for both massed and distributed groups. Thus, it appears that off-target movements are too slow and too short in radius. The effect of massed practice was to exaggerate these errors; the difference between following versus leading, and inside versus outside, was greater under massed than under distributed practice. Thus, accumulated reactive inhibition resulted in poorer performance presumably by slowing and by radially shortening the rotary pursuit-movements. Therefore it is not unreasonable to think that both reactive inhibition and satiation produce 'distortion' along certain behavorial dimensions.

(4) *Time to accumulate in measurable amounts.* Although we have concluded that both processes develop from afferent stimulation, measurable effects of both require more than momentary stimulation. It is possible to obtain improvement in performance on a task without any overt evidence of I_R if practice is sufficiently distributed. In perception, both inspection- and test-figures will be perceived without distortion if each is presented only briefly. The question then is: How long must stimulation continue before each process begins to exhibit its typical effects?

Both Gibson and Radner, and Hammer, collected systematic data on this point and both found visual figural after-effects after 5 sec. of continuous stimulation.[19] Köhler and Wallach agree that 5 sec. may be sufficient with some figures.[20] The latter investigators also report that satiation will accumulate in amounts sufficient to produce figural after-effects even if stimulation is not continuous but is rapidly repetitive. In one case they exposed the inspection-figure tachistoscopically for 0.25 sec., with 1.25 sec. between exposures, and found figural after-effects after as few as 20 exposures, or a total of 5 sec. exposure-time.

The effects of reactive inhibition are also observable after fairly brief periods of either continuous or highly massed practice on motor tasks. Unfortunately, there has been no systematic study of periods of stimulation shorter than 60 sec. Ammons found that about 35% of the maximal amount of reminiscence occurred after 20 sec. of continuous practice on the rotor.[21] Kimble and Bilodeau found clearly poorer performance on the first trial of block turning when the trial-duration was 30 sec. than when it was 10 sec.[22] With reversed alphabet-printing, Wasserman found 80-100% of the maximal amount of reminiscence after 60 sec. of practice.[23] Bilodeau varied weight of balls rather than work-time in a task involving picking up balls and dropping them through a chute as rapidly as possible.[24] He found a

[19] Gibson and Radner, *op. cit.,* 453-467; E. R. Hammer, Temporal factors in figural after-effects, this JOURNAL, 62, 1949, 337-354.
[20] Köhler and Wallach, *op. cit.,* 325.
[21] Ammons, *op. cit., J. exp. Psychol.,* 37, 1947, 393-411.
[22] Kimble and Bilodeau, Work and rest as variables in cyclical mootr learning, *J. exp. Psychol.,* 39, 1949, 150-157.
[23] H. N. Wasserman, The effect of motivation and amount of pre-rest practice upon inhibitory potential in motor learning, *J. exp. Psychol.,* 42, 1951, 162-172.
[24] Bilodeau, *op. cit.,* 381-390.

significant difference in performance between loads of 21 gm. and 168 gm. within the first 5 sec. of work. Denny, Frisbey, and Weaver found that when they analyzed a 30-sec. trial on the rotor into successive 10-sec. thirds, performance on the last third was always poorer than on the first.[25] Thus, even though the data are not as systematic for I_R as for satiation, it can probably be concluded that with continuous stimulation the effects of both processes will be manifest after periods of stimulation as brief as 5–10 sec.

(5) Time to produce most intense effects. With continuous stimulation both satiation and I_R seem to develop rather quickly to a kind of maximum, or possibly an equilibrium between accumulation and dissipation, such that additional stimulation does not further intensify the behavioral disortions.

Both Gibson and Radner and Hammer report systematic data showing that after 60 sec. of inspection-time visual figural after-effects were as severe as they would ever be.[26] Köhler and Wallach also say that prolonged inspection-times do not increase the magnitude of after-effects.[27] In the case of I_R, Ammons found that about 54% of the maximal reminiscence on the rotor occurred after 60 sec. of practcice, and 78% after 3 min.[28] Irion found very nearly the maximal amount of reminiscence after 250 sec. of rotor practice in which nine 5-sec. rest periods were interspersed.[29] Wasserman reported that highly motivated Ss yielded their maximal reminiscence, while less highly motivated Ss yielded 80% of their maximum, after only 60 sec. of practice.[30] In all of these cases additional massed practice was given and could have yielded considerably more reminiscence (since performance was well below the asymptote), but did not.

The characteristic depression of performance under massed practice also indicates that reactive inhibition rises quickly to some maximum after which its effect on performance does not worsen. Studies that present performance-curves for both a massed and a spaced group during trials before a long rest is given the massed group (for example, Adams and Reynolds, Duncan, Kimble, Kimble and Shatel, and Wasserman show that the massed-practice curve falls below the distributed curve very quickly after the start of practice and then tends to remain fairly level or to increase slightly.[31] Even on the 'pure work' tasks used by Bilodeau, where the

---

[25] M. R. Denny, Norman Frisbey, and John Weaver, Jr., Rotary pursuit performance under alternate conditions of distributed and massed practice, *J. exp. Psychol.*, 49, 1955, 48-54.

[26] Gibson and Radner, *op. cit.*, 467; Hammer, *op. cit.*, 354.

[27] Köhler and Wallach, *op. cit.*, 269-357.

[28] Ammons, *op. cit.*, *J. exp. Psychol.*, 37, 1947, 393-411.

[29] Irion, Reminiscence in pursuit-rotor learning as a function of rest and of amount of pre-rest practice, *J. exp. Psychol.*, 39, 1949, 492-499.

[30] Wasserman, *op. cit.*, 162-172.

[31] J. A. Adams and Bradley Reynolds, Effect of shift in distribution of practice conditions following interpolated rest, *J. exp. Psychol.*, 47, 1954, 32-36; C. P. Duncan, The effect of unequal amounts of practice on motor learning before and after rest, *ibid.*, 42, 1951, 257-264; Kimble, *op. cit., ibid.*, 39, 1949, 15-23; Evidence for the role of motivation in determining the amount of reminiscence in pursuit rotor learning, *ibid.*, 40, 1950, 248-253; G. A. Kimble and R. B. Shatel, The relationship between two kinds of inhibition and the amount of practice, *ibid.*, 44, 1952, 355-359; Wasserman, *op. cit.*, 162-172.

compensating effect of increasing habit-strength presumably does not enter, performance-curves typically drop sharply at first and then level off.[32] Thus, with continued stimulation there is a point reached beyond which the distortions produced by either satiation or reactive inhibition do not further increase.

(6) The rate of immediate decay. Both satiation and I_R begin to dissipate very rapidly immedaitely after cessation of stimulation.

With reference to I_R, Ammons, Bilodeau, Grice and Reynolds, Irion, and Kimble and Horenstein measured reminiscence-gains over rest (following masse dpractice) at several points in the region running in general from 10 sec. to 10 min. and all found that the amount of reminiscence increased sharply during the first few minutes.[33] The findings for satiation are similar. Although they do not present data, Köhler and Wallach noted that visual after-effects begin to decrease rapidly after removal of the inspection-figure.[34] Bales and Follansbee, Hammer, and Krauskopf, have reported data for various intervals ranging from 0 to 180 sec. and all found that the magnitude of figural after-effects decreases rapidly after removal of the inspection-figure.[35] Thus, both satiation and reactive inhibition dissipate over rest and the rates of decay for both are quite rapid at first.

THE CONFLICTING DATA ON THE PERSISTENCE OF EACH PROCESS

Comparison of the time required for the complete dissipation of satiation and of I_R following a single stimulation-period is difficult because for each process the data are not consistent. Until recently it did appear that studies of the persistence of reactive inhibition agreed closely. Using reminiscence-gain over rest as the measure of decay of I_R, Ammons, Bilodeau, Ellis, Montgomery, and Underwood, Grice and Reynolds, and Kimble and Horenstein all reported an apparent asymptote in the curve of reminiscence-gain after 5–10 min. of rest.[36] In fact, since Ammons found no further reminiscence-gain in the interval from 10 min. to 6 hr., apparently no one has questioned the assumption that all the reactive inhibition that is generated by a few minutes of massed practice is dissipated within 10 min. or so of rest. Yet it now appears that this assumption may be false. As part of a larger study now being conducted, the reminiscence-gain over rests of 10 min., 24 hr., and 7 days with independent groups of Ss that had been given 6 min. of massed practice on the pursuit-rotor, is now being tested. The results are clear-cut; there was slightly more reminiscence after 24 hr. than after 10 min., and more after 7 days

[32] Bilodeau, *op. cit.*, 381-390; 96-100.

[33] Ammons, *op. cit., J. exp. Psychol.*, 37, 1947, 393-411; 43, 1952, 381-390; Grice and Reynolds, *op. cit.*, 247-252; Irion, *op. cit.*, 492-499; G. A. Kimble and B. R. Horenstein, Reminiscence in motor learning as a function of length of interpolated rest, *J. exp. Psychol.*, 38, 1948, 239-244.

[34] Köhler and Wallach, *op. cit.*, 269-357.

[35] Bales and Follansbee, *op. cit.*, 499-503; Hammer, *op. cit.*, 337-354; Krauskopf, The magnitude of figural after-effects as a function of the duration of the test-period, this JOURNAL, 67, 1954, 684-690.

[36] Ammons, *op. cit., J. exp. Psychol.*, 37, 1947, 393-411; Bilodeau, *op. cit.*, 381-390; D. S. Ellis, Victor Montgomery, and B. J. Underwood, Reminiscence in a manipulative task as a function of work-surface height, prerest practice, and interpolated rest, *J. exp. Psychol.*, 44, 1952, 420-427; Grice and Reynolds, *op. cit.*, 247-252; Kimble and Horenstein, *op. cit.*, 239-244.

than after 24 hr.[37] Since it was also ascertained that following distributed practice the rotor habit shows some forgetting over these intervals, it is clear that the reminiscence-gain in the massed groups more than counteracted forgetting. Thus, it is possible that even though reactive inhibition dissipates very rapidly in the first few minutes after cessation of stimulation, traces of the process may remain for much longer periods.

The data on satiation are also somewhat conflicting. Gibson had one S who reported after-effects to persist well into the second day after four days of wearing prisms that made vertical lines appear curved.[38] Bales and Follansbee, using a 10-min. inspection-period, reported as great an after-effect one week later as they had found on the immediate test.[39] Köhler and Wallach say that some after-effects may not disappear for hours while Köhler and Dinnerstein, studying kinesthetic after-effects, found that satiation increased regularly for more than a month when they gave a single 60-sec. stimulation-period each day.[40] The longest intervals have been reported by Köhler and Fishback.[41] They summarize several cases where visual and kinesthetic after-effects were still measurable after intervals ranging up to four and five months. (Although many of these studies are based on use of the same Ss in successive tests, it is unlikely that this would completely negate the findings.)

In contrast to these reports, we have the careful quantitative study by Hammer.[42] She took readings at several points in the intervals from 0 to 180 sec. following stimulation and found no measurable after-effect after 90 sec. (Successive tests with the same Ss gave the same results as tests with independent groups.) The discrepancy between this finding and those cited above is not easily resolved, but two suggestions may be offered. First, there may be differences due to the inspection- and test-figures used. Köhler and Wallach state that different inspection-times are often required for different figures. A more important point is the length of the inspection-time. Hammer used a 60-sec. inspection-interval, whereas all of the studies that have found long-persisting after-effects used intervals of 2-10 min. or even more. Köhler and Wallach stated specifically their belief that although the intensity or magnitude of after-effects does not increase indefinitely with increasing inspection-times—see (5) above—continued inspection makes for more persistent after-effects. Gibson and Radner had earlier reported the same impression.[43]

---

[37] These findings raise some doubts about theories of motor performance which assume that some additional process, such as conditioned inhibition, is needed to account for any residual depression of performance which remains in curves of massed practice after about 10-min. rest. At least a part of what has been called conditioned inhibition may be merely residual reactive inhibition. The findings may also provide an explanation for the fact that when massed practice is resumed after several minutes of rest, performance curves sometimes actually decrease as practice continues, as shown by Kimble and Duncan in 1950. Reactive inhibition should reach its maximum quickly when residual amounts are still present, and at this point in practice the compensating effects, i.e. increments to habit strength, are smaller than at the start of practice. These points will be discussed more thoroughly when the complete study is reported.

[38] Gibson, op. cit., J. exp. Psychol., 16, 1933, 1-31.
[39] Bales and Follansbee, op. cit., 499-503.
[40] Köhler and Dinnerstein, op. cit., 196-220.
[41] Köhler and Fishback, op. cit., 398-410.
[42] Hammer, op. cit., 337-354.
[43] Gibson and Radner, op. cit., 453-467.

Thus, it is possible that after-effects will show great persistence only after inspection-times of several minutes' duration.

Since the data on persistence of either reactive inhibition or neural satiation are not in agreement, it is obvious that we do not have good grounds for claiming that the two processes are similar in this characteristic. Neither would we be justified, however, in concluding that we have located a basic difference between them. Instead we shall offer the following merely as an hypothesis: At the beginning of a rest following several minutes of continuous stimulation, such as 5 min. of continuous practice on a motor task or 5 min. inspection of a visual figure, both reactive inhibition and satiation will, as usual, begin to decrease very rapidly. Most of the total amounts built up will dissipate within the first few minutes. Residual effects may, however, be detected over much longer periods of time. In short, we shall assume that satiation and I_R have similar decay functions over all periods of time and that under certain conditions the time for complete decay may be quite long.

If future research supports the hypothesis that the rate and asymptote of decay are the same for both processes, it would appear that they are not dissimilar in any important characteristic. There would then be some grounds for concluding that we have two names for the same basic process.

SUMMARY

Some of the literature dealing with performance on motor tasks and some dealing with figural after-effects has been reviewed that the processes of reactive inhibition and neural satiation might be compared. It appeared that the two processes are quite similar on most, perhaps all, characteristics. Although more research is needed on some points, present evidence suggests that reactive inhibition and neural satiation may refer to the same process.

DISCRIMINATIVE AND VERBAL HABITS IN
INCIDENTAL LEARNING

By MAVIS PLENDERLEITH and LEO POSTMAN, University of California

This paper is concerned with discriminative and verbal habits that are significantly related to variations in the degree of incidental learning. Under constant conditions of practice, incidental learners typically show a high degree of variation in amount retained. Frequently individual differences under incidental conditions are larger than under comparable conditions of intentional learning.[1] Some of the correlates of these individual differences are explored in the present experiment.

Analysis of incidental learning under a variety of experimental conditions suggests that there are two characteristics of Ss which should be significantly related to success in incidental learning. First, we should expect degree of incidental learning to vary with S's ability to maintain a *multiple set, i.e.* to discriminate and categorize stimulus-materials along more than one dimension. In experiments on incidental learning, S must perform some orienting task which insures his exposure to the stimulus-materials. The orienting task focuses S's set on manipulations of the materials which, by definition, must be different from those involved in intentional practice. Degree of incidental learning should, then, vary with S's disposition to generalize his set from the orienting task to the differentiation of the stimulus-items. A second determinant of success in incidental learning which is at once suggested is the *availability and effectiveness of differential responses* to the stimulus-items. Recall depends on the degree to which the stimulus-items have been differentiated during exposure. The more effective are S's differential responses to the learning stimuli, the more probable it is that the items can be reproduced at the time of recall.[2] In the case of verbal stimuli, differential responses are themselves likely to be verbal or symbolic in nature. The meaningful asso-

---

[1] For an example of such differences in variability, see Leo Postman and L. W. Phillips, Studies in incidental learning: I. The effects of crowding and isolation, *J. exp. Psychol.,* 48, 1954, 48-56.

[2] The role of differential response in incidental learning has been discussed by Leo Postman, P. A. Adams and L. W. Phillips (Studies in incidental learning: II. The effect of association value and of the method of testing, *J. exp. Psychol.,* 49, 1955, 1-10). See also George Mandler, Response factors in human learning, *Psychol. Rev.,* 61, 1954, 235-244.

ciations evoked by nonsense-syllables, which play a critical role in the memorization of such materials, provide an excellent example. In general, differential responses to verbal items include classifying or categorizing responses such as naming or labeling, and other associative responses. Such responses sample the verbal associative patterns of the individual, both as they derive from the structure of the language and as they grow out of his individual history. The more readily available such differential responses are to an incidental learner, the better should be his recall for verbal materials.

The factors which we have considered critical for incidental learning are clearly relevant to intentional learning as well. Differentiation of the stimulus-items is as essential in intentional as in incidental learning. Ability to maintain a multiple discriminative set should be related to success in intentional learning to the extent that the material is complex and requires differentiation along several dimensions. Incidental learning should, however, be more sensitive than intentional learning to differences in these verbal and discriminative skills. Under incidental conditions, Ss are not motivated to use their skills for purposes of differentiating and connecting the stimulus-items. Since only differential responses which are strong and readily available will be evoked by the stimuli, individual differences in the repertoire of such responses become critical. Under intentional conditions, Ss are motivated to utilize maximally whatever skills they have in order to differentiate and integrate the stimulus-items. As long as the learning task remains simple, high motivation can compensate for limited skill and individual differences may remain masked. As the learning task grows in complexity, however, individual differences in discriminative and verbal skills should be reflected increasingly in the performance of the intentional Ss. These considerations suggest the following hypotheses: (a) degree of incidental learning should correlate more highly than degree of intentional learning with measures of the discriminative and verbal skills described above; (b) the correlation between degree of intentional learning and measures of discriminative and verbal skills should increase as the learning task grows in complexity; (c) the correlation between incidental and intentional learning should vary with the complexity of the intentional learning task. As that task grows in complexity, the correlation between intentional and incidental learning should increase.

METHOD

The experiment comprised four procedures: (a) incidental learning, (b) intentional learning, (c) symbol-discrimination under single and multiple sets, and (d)

solution of anagrams, a performance which was used to gauge the effectiveness of differential responses to nonsense-stimuli.

Incidental learning. The stimulus-materials were a list of 20 nonsense-syllables. There were 5 syllables each of 0, 33, 66, and 100% association value by Glaze's norms.[3] The series was divided into five blocks, with each block containing one syllable of each association-value. (Previous experimentation had shown that the incidental learning of nonsense-syllables is facilitated by the use of items varying widely in associative value.[4]) The orienting task required S to match the syllables with geometric designs. For each syllable, S selected one of the following geometric designs as best fitting the item: square, triangle, circle, cross, arrow, diamond, wavy line (sine curve). The instruction was to make quick 'intuitive' matches. Each syllable was exposed for 4 sec. on a 2 × 2 in. slide. Following the exposure, S was given 6 sec. in which to draw the chosen figure in an appropriately numbered space of a record-blank. A cardboard strip was used to cover up all spaces except the one that S was writing in. Large-scale copies of the designs were kept in full view throughout the procedure. After one exposure of the series, retention of the nonsense-syllables was tested by the method of free recall. The test had a time-limit of 5 min., following which S was asked whether he had attempted to learn the syllables or had expected a test of recall. Only two Ss gave an indication that they had guessed the purpose of the experiment and they were replaced.

Intentional learning. The stimulus-series consisted of 20 nonsense-syllables, with an average association value of 25%. None of the items had more than one letter in common with the syllables used for incidental learning. The syllables, printed in capital letters (2 in. high) on 4 × 7 in. flash-cards, were exposed at the rate of one every 2 sec. The series was presented four times, each time in a different random order, and S was instructed to try to remember as many items as possible, regardless of order of presentation. Following each presentation of the list, S was given 3 min. in which to write down all the items he could remember. He was encouraged to guess if in doubt about any syllable.

Symbol-discrimination. The discriminative task involved a procedure similar to that used in the measurement of the span of apprehension. A series of slides was exposed tachistoscopically. The slides were projected on a milk-glass screen by means of an *SVE* projector, and a photographic shutter, attached to the lens of the projector, was used to regulate speed of exposure. The duration of exposure was kept at 0.60 sec. throughout. A pattern of 12 symbols—varying numbers of capital letters, small letters, and digits—appeared on each slide, with the spatial arrangement of the three kinds of symbol varying from slide to slide. For the first 8 slides in the series, S was instructed to discriminate and report on digits only, *i.e.* he operated under a single set. The numbers seen were recorded by S in the appropriately numbered space of a record-blank. For the remaining 13 slides of the series, S was instructed to discriminate both digits and capital letters, but not told until after the exposure of each slide which of the symbols to report on, *i.e.* he operated under a multiple (double) set. Reports on digits were obtained for 8 (of the 13) slides which were matched with those used in the first part of the series with

---

[3] J. A. Glaze, The association value of non-sense syllables, *J. genet. Psychol.*, 35, 1928, 255-267.

[4] Postman, Adams, and Phillips, *op. cit.*, 3-5.

respect to the number of digits appearing in the patterns of symbols. Reports on capital letters were obtained for the other 5 slides. Reports on letters were interspersed at random among the reports on digits in order to maintain S's multiple set. The decrement in accuracy of discimination of digits from the first to the second part of the series provided a measure of S's ability to maintain a multiple set. The larger the decrement, the less skilled was the S in the simultaneous discrimination of different features of the stimulus-materials.

Solution of anagrams. The anagram-task was adapted from a test used by Lowell.[5] It consisted of 10 pages containing 24 anagrams each. The anagrams were formed from words which appear among the 500 most frequent items in the Thorndike-Lorge word-counts.[6] The time allowed was 2 min. per page, which was insufficient for the solution of all the items. The anagram-task provided a measure of S's ability to differentiate nonsense-stimuli in terms of their similarity to conventional meaningful units.

Order of tasks. The four procedures occupied two experimental sessions separated by one week. The first session always began with incidental learning, and the second session always began with intentional learning. This order (*a*) prevented the transfer of learning sets from the other tasks to the incidental procedure, and (*b*) kept the interval between the two learning tasks constant for all Ss. For half the Ss, symbol-discrimination occurred during the first session and the anagram-task during the second session. For the other half of the Ss, the order of symbol-discrimination and anagram-solution was reversed.

Subjects. The Ss were 100 undergradaute students. They were tested in groups averaging about five in number. They did not know the purpose of the experiment.

RESULTS

Group averages. Table I shows the average scores of the total group for the four experimental tasks. After one trial of intentional learning the level of retention was considerably higher than after one trial under incidental conditions.[7] The intentional learners, of course, improved steadily from trial to trial in the number of items recalled. After four presentations, the level of recall was still far from perfect (66.15%). In the symbol-discrimination task, the average number of digits correctly recognized was significantly greater with a single than with a multiple set (t = 8.88, P < 0.01). The drop was from 39.83% to 33.42% correct recognitions, and it occurred in spite of the fact that the first part of the series provided opportunity for practice in the perceptual task. The average number of anagrams solved was 111.5, out of a total possible of 240. The task was, therefore,

---

[5] E. L. Lowell, The effect of need for achievement on learning and speed of performance, *J. Psychol.*, 33, 1952, 31-40.

[6] E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words,* 1944.

[7] This difference was obtained in spite of the fact that the average association value of the syllables was appreciably lower in the list learned intentionally than in the list learned incidentally.

difficult enough to allow individual differences in skill to manifest themselves.

Correlates of the degree of incidental learning. Table II presents the correlations between degree of incidental learning and the other experimental measures. At this point we consider only the first trial of intentional learning which is most comparable to incidental learning with respect to

TABLE I

MEAN SCORES FOR THE VARIOUS EXPERIMENTAL TASKS

Task	Mean	SD
Incidental learning	2.71	1.54
Intentional learning		
Trial 1	4.55	2.10
Trial 2	8.00	3.01
Trial 3	11.24	3.35
Trial 4	13.23	3.64
Symbol-discrimination		
Single set	14.34	3.65
Multiple set	12.03	3.01
Anagrams	111.05	29.18

the amount of training. The symbol-discrimination task is represented by a measure of 'discrimination-decrement'—the difference between the number of items correctly recognized under single and multiple set. The use of a difference-score makes measures of the effect of set independent of absolute level of acuity; the smaller the decrement, the better S is able to

TABLE II

INTERCORRELATIONS OF INCIDENTAL LEARNING, INTENTIONAL LEARNING (FIRST TRIAL), DISCRIMINATION-DECREMENT, AND ANAGRAM-SOLUTION

	Intentional learning	Discrimination-decrement	Anagrams
Incidental learning	.26*	−.42*	.25*
Intentional learning		.16	.11
Discrimination-decrement			−.21†

* Significant at the 1% level of confidence.
† Significant at the 5% level of confidence.

function under a multiple set. Performance on the anagram-task is represented by total number of correct solutions.

Degree of incidental learning correlates more highly than does degree of intentional learning with both discrimination-decrement and performance on the anagram-task. Both correlations with incidental learning are significant at the 1% level of confidence, although the relationship with

discrimination-decrement is clearly more pronounced. Neither of the correlations with intentional learning is significant. The difference in the relationship with intentional and incidental learning is particularly striking in the case of discrimination-decrement; in the case of anagrams, the difference is smaller but in the same direction. The negative correlation between discrimination-decrement and performance on the anagram-task suggests that the discriminative and verbal skills measured by these two procedures are not entirely independent of each other. It appears that, in the early stages at least, incidental learning is more sensitive to individual differences in verbal and discriminative skills than is intentional learning. This conclusion is supported by the fact that the correlation between degree of incidental learning and performance on the first trial of intentional learning is relatively low although significant.

Correlations at different stages of intentional learning. We have suggested earlier that with a relatively simple learning task the high motivation in intentional learning will serve to mask some of the effects of the individual differences to which incidental learning is sensitive. As the intentional learning task becomes more complex, we expect the two types of learning to show increasingly similar patterns of correlation with measures of discriminative and verbal skills.

As intentional learning progresses from trial to trial, S's task may be assumed to grow in complexity. An increasing number of items must be differentiated and integrated. The pattern of associative connection becomes more complex and there is increasing opportunity for intra-serial interference. Continuing progress in learning depends, therefore, on a higher order of associative habits than is required in the earlier stages of practice, and the successful S must make maximal use of meaningful associations and other mnemonic devices. These considerations suggest that the habits and skills favoring incidental learning should gain increasing importance at successive stages of intentional learning. We expect, therefore, that performance on successive trials of intentional learning will correlate more and more highly both with the measures of discriminative and verbal skills and with incidental learning. Table III shows that such is indeed the case.

For both discriminative efficiency and solution of anagrams, the correlations with intentional learning show a systematic increase as a function of the number of trials. There are small reversals on the last trial. The increase cannot, of course, be expected to continue indefinitely. As more and more Ss approach high levels of recall, correlations with individual differences must necessarily decrease. It is true that the correlations with

intentional learning are relatively low and remain below those with incidental learning. The point to be emphasized is the systemic trend in the correlations which serves to make the general pattern of relationships increasingly similar for intentional and incidental learning. At the same time, the fact that the correlations with intentional learning remain lower than those with incidental learning suggests that the higher motivation in the intentional situation continues to mask the effects of individual differences in the skills measured.

The correlation between intentional and incidental learning rises steadily

TABLE III

CORRELATION OF SCORES ON SUCCESSIVE TRIALS OF INTENTIONAL LEARNING WITH
THE OTHER EXPERIMENTAL MEASURES

Procedure	Trial			
	1	2	3	4
Incidental learning	.26*	.36*	.52*	.56*
Discrimination-decrement	.16	−.01	−.24†	−.18
Anagrams	.11	.12	.24†	.19†

* Significant at the 1% level of confidence.
† Significant at the 5% level of confidence.

on successive trials of intentional learning. The increase over the four trials is 30 correlation points, although the trend is clearly leveling off after the third trial. This rise fits the general conclusion that the abilities critical for incidental learning become increasingly important as intentional learning progresses. The increasing covariation of intentional and incidental learning cannot, however, be fully accounted for in terms of the variables measured by symbol-discrimination and solution of anagrams since the correlations between these variables and intentional learning remain rather low. Undoubtedly there are verbal and discriminative skills essential for both intentional and incidental learning which are not measured by our tasks. The trend in the correlations strongly supports the conclusion, however, that the same factors which favor incidental learning gain in importance as the products of intentional learning grow in size and complexity.

Multiple correlations with incidental learning. How well can individual differences in incidental learning be predicted on the basis of performance in intentional learning, symbol-discrimination, and solution of anagrams? Since the pattern of intercorrelations changes as a function of the number of intentional learning trials, the overall efficiency of prediction also varies with the stage of intentional learning. Separate coefficients of multiple

correlation were computed using each of the four trials of intentional learning in turn. The resulting values of R were: 0.54, 0.56, 0.61, 0.65.[8] These coefficients are of moderate magnitude. Important sources of individual differences in incidental learning remain to be explored.

SUMMARY

This experiment was designed to investigate the discriminative and verbal habits related to individual differences in incidental learning. Ability to maintain a multiple set (measured by symbol-discrimination) and effectiveness of differential response to verbal stimuli (measured by anagram-solution) were correlated with measures of incidental learning and intentional learning. Analysis of the pattern of intercorrelations yielded the following results: Degree of incidental learning correlates more highly than degree of intentional learning with proficiency in the discriminative and verbal tasks. Incidental learning is more sensitive than intentional learning to individual differences in verbal and discriminative skill. The later trials of intentional learning correlate more highly with performance in the verbal and discriminative tasks than do the earlier trials. Successive trials of intentional learning also show a steady rise in correlation with incidental learning. The factors which favor incidental learning also gain in importance as the products of intentional learning gain in number and complexity.

---

[8] In computing the coefficients of multiple correlation we used the method described by J. A. Gengerelli (A simplified method for approximating multiple regression coefficients, *Psychometrika*, 13, 1948, 135-146) which provides a minimal estimate of the value of the coefficient.

EXTINCTION OF A HUMAN CARDIAC-RESPONSE DURING AVOIDANCE-CONDITIONING

By P. J. Bersh, Human Factors Laboratory, Rome Air Development Center, Rome, N.Y., J. M. Notterman and W. N. Schoenfeld, Columbia University

Since avoidance-conditioning always involves, at least at the outset, the exposure of S to the negatively reinforcing ('noxious' or 'to-be-avoided') stimulus, it is now generally recognized that the Pavlovian paradigm, with the negatively reinforcing stimulus acting as US, applies with significant consequences. The joint occurrence of autonomically-mediated CRs and the instrumental avoidance-response (R_{AV}) raises the possibility of interaction between the two response-systems,[1] and this possibility makes desirable the simultaneous measurement and study of representative reflexes in both systems.

One initial hypothesis on which agreement might presently be expected is that the autonomic CRs acquired during avoidance-training tend toward extinction on 'successful' avoidance-trials (*i.e.* trials on which the experimentally prescribed R_{AV} is made and US is not forthcoming), in accordance with the Pavlovian paradigm for extinction. Nevertheless, the literature seems to contain no data indicating that such extinction actually occurs, and one purpose of the experiments reported here is to furnish that evidence. It will also appear that the weakening of CR arises partly from the discriminative control over CR assumed by stimuli which are produced by, or accompany, R_{AV}, and which thus play the role of negative stimuli in the Pavlovian method of contrasts.[2]

[1] W. N. Schoenfeld and J. J. Antonitis, A function of respondents in the extinction of operant responses, *Notes, Conf. Exper. Anal. Behav.*, No. 17, May 9, 1949, issued from Indiana Univ. (mimeographed); Schoenfeld, An experimental approach to anxiety, escape, and avoidance behavior, in *Anxiety*, edited by P. H. Hoch and Joseph Zubin, 1950, 70-99; J. M. Notterman, Experimental anxiety and a conditioned heart rate response in human beings, *Trans. N. Y. Acad. Sci.*, Ser. II, 16, 1953, 24-33.

[2] W. N. Schoenfeld, P. J. Bersh, and J. M. Notterman, Interaction of instrumental and autonomic responses in avoidance conditioning, *Science*, 120, 1954, 788, (Abstract).

EXPERIMENT I

The purpose of the experiment was to compare the magnitude of a cardiac *CR* after acquisition of a motor avoidance-response with *CR*-magnitude prior to avoidance-training.

Procedure. In the first experiment, as well as in two others to be reported, the autonomic *CR* under observation was one previously studied by the authors as a function of several parameters, *i.e.* the decrease in heart-rate produced in human *S*s by repeated exposure to tone (*CS*) followed by electric shock (*US*).[3] The same laboratory arrangements, apparatus, and general procedure were employed as in the previous studies cited. The cardiac *CR* was measured by comparing pre-*CS* heart-rate with rate during the *CS–US* interval, and a trace-conditioning procedure was used, consisting of a 1-sec. tone, a 6-sec. tone-shock interval, and a 6-sec. shock. Additional particulars of procedure are given for each of the three experiments reported.

Twenty college students (men) served as *S*s.[4] There were three experimental phases: (a) determination of basal (pre-conditioning) heart-rate response to the tone, (b) cardiac conditioning, (c) establishment of an avoidance-response (key-tapping). All phases were administered to each *S* in a single session lasting approximately 2 hr.

After *S* entered the laboratory, he was asked to strip to the waist, and the cardiograph and shock electrodes were placed. During these preparations, the following instructions were given by E.

This is an experiment to determine your heart-rate reaction to tone, and your heart-rate reaction to electric shock. You will hear a series of tones through this headset which will soon be placed over your ears. A little later in the experiment you will receive a series of electric shocks administered to your left hand; you will not be shocked through the electrodes I am now placing on your chest and left leg. These electrodes enable me to record your heart action. The shocks you will receive will be strong enough for you to feel them, but will not be strong enough to do you any harm. Remember that they will just pass through your hand, and not your body. Please try not to make any unnecessary movements. I will be keeping a continuous record of your heart-action while you are in here, and since the apparatus is necessarily very sensitive, your moving around too much would interfere with the recording.[5]

---

[3] Notterman, Schoenfeld, and Bersh, Conditioned heart rate response in human beings during experimental anxiety, *J. comp. physiol. Psychol.*, 45, 1952, 1-8; Partial reinforcement and conditioned heart rate response in human subjects, *Science*, 115, 1952, 77-79; A comparison of three extinction procedures following heart rate conditioning, *J. abnorm. soc. Psychol.*, 47, 1952, 674-677; Bersh, Schoenfeld, and Notterman, The effect upon heart rate conditioning of randomly varying the interval between conditioned and unconditioned stimuli, *Proc. Nat. Acad. Sci.*, 39, 1953, 563-570.

[4] These *S*s remained of an original group of 28. Three *S*s were discarded for failure to show the cardiac *CR* at all within the 8 trials set aside for acquisition (compare number of trials in earlier studies cited). Five *S*s were discarded for failure to learn R_{AV} to the given criterion (see "Procedure") before 17 shocks were administered. These shocks, together with those of acquisition, represented about the maximal number that *S*s cared to endure during a single 2-hr. session.

[5] *Cf.* the procedure and instructions of earlier studies cited.

Then E left S's room and entered the adjacent apparatus room. The basal trials were begun immediately thereafter. A series of 20 1-sec. tones, spaced irregularly 1–2 min. apart, was sent through the headset, S's cardiogram being taken continuously from about 30 sec. before to about 30 sec. after each tone. As in the previous studies, rate was measured by determining the period of the last two cardiac cycles immediately preceding onset of tone, and the period of the last two cycles in the 6-sec. interval following tone (and, therefore, immediately preceding shock on the subsequent tone-shock trials). These measures, called the 'pre-tone' and 'post-tone' rates, respectively, were then converted to beats-per-min.

Following the basal series, cardiac conditioning was begun immediately for all Ss without further instruction. This phase consisted of eight tone-shock trials. It may be noted here that the first occasion on which tone is followed by shock (Trial 1 of cardiac conditioning) is, insofar as S's cardiac response to tone is concerned, another basal trial. It is only on the second trial of conditioning that any conditioning effect to tone may be exhibited. For this reason, in the computations and tables reported, the first trial of cardiac conditioning has been treated as the last basal trial.

Upon conclusion of the eighth cardiac conditioning trial, all Ss were given the following instructions.

From now on, if you tap this key once as soon as you hear the tone, and once again at a certain specific time-interval after the first tap, you will not receive the shock. You have to determine for yourself what the proper interval is between the two taps. Keep your hand near the key.

These instructions were repeated a second time, after which a telegraph key was put in position on the table near S's right hand. If, on the succeeding motor-conditioning, S tapped the key for the second time during the fourth second after termination of CS, shock was not given at the conclusion of the usual 6-sec. trace-period. Motor-conditioning was considered achieved when S successfully avoided shock on five consecutive trials (the criterion-trials) before receiving 17 shocks. During each trial of the motor-conditioning phase, heart-rate was recorded in the usual way (from 30 sec. before tone to 30 sec. after shock would normally terminate).

Results. The cardiac measures are summarized in Table I. The difference between the mean pre-tone and post-tone rates during the basal phase is not significant.[6] The effect of the conditioning trials is shown in the conditioned decrement which is significantly greater than zero and also significantly greater than the basal decrement ($P < 0.01$). During R_{AV} criterion-trials, the magnitude of CR is reduced; it is still significantly above zero and above the basal drop ($P < 0.05$), but significantly smaller than the CR during acquisition ($P < 0.05$). The weakening of CR during the R_{AV} criterion-trials is further attested by the fact that 15 of the 20 Ss (15/20 gives $P < 0.03$) show a smaller CR during the R_{AV} criterion-trials than they did at the

---

[6] All conclusions concerning statistical significance are based on Wilcoxon's nonparametric tests for paired and non-paired replicates (Frank Wilcoxon, Individual comparisons by ranking methods, *Biometric Bull.*, 1, 1945, 80-83).

end of acquisition (despite the increase in resistance to extinction which presumably arises from the partial reinforcement schedule to which CR is exposed during the acquisition of R_{AV}).[7] In brief, the conclusion indicated by the data is that autonomically-mediated CRs involved in avoidance training do, in fact, tend to extinguish in the expected manner during R_{AV} trials.

One possible artifact to be considered in connection with this conclusion is the unconditioned effect of R_{AV} (in this case, the tapping of a telegraph key) upon the cardiac-measure employed, since it is commonly recognized that skeletal responses tend to raise cardiac-rate. This effect, counteracting the cardiac CR-drop, might have created the false impression that extinction had occurred. That this artifact did not operate in the present experiment is indicated by control data reported elsewhere.[8]

While the present data indicate that the effect of R_{AV} on CR is in the expected direction, this effect is not large, since CR remains significantly above zero and above the basal level. Although many conditions of the experiment might account for such a result, including the particular R_{AV} criterion involved, one likely basis may be the 'externally unsupported' nature of the temporally-discriminated R_{AV} used here. The possibility of speeding extinction of the CR by providing an exteroceptive stimulus associated with, and marking, R_{AV} therefore suggests itself. This possibility led to the performance of Experiment II.

EXPERIMENT II

The purpose of this experiment was to determine whether an exteroceptive stimulus associated with an avoidance-response influences the extinction of the associated cardiac CR.

Procedure. Ten students (men) served as Ss.[9] The procedure of this experiment was identical with that of Experiment I, except that, immediately upon occurrence of a 'successful' avoidance-response, E presented a signal-light of 1.5-sec. duration. The light was provided by a small lamp positioned on the wall in front of S and approximately at eye-level. On trials during which S made an 'incorrect' key-tapping response, no light-signal was given. No mention was made of the light-stimulus during the instructions or during any phase of the experiment.

---

[7] Notterman, Schoenfeld, and Bersh, Partial reinforcement and conditioned heart rate response in human subjects, *op. cit.,* 77-79.

[8] Notterman, Experimental anxiety and a conditioned heart rate response in human beings, *op. cit.,* 31; Notterman, Schoenfeld, and Bersh, A comparison of three extinction procedures following heart rate conditioning, *op. cit.,* 676-677.

[9] Two other Ss were eliminated from the experiment for failure to demonstrate cardiac conditioning, and six for failure to attain the criterion of avoidance-conditioning criterion, within the number of trials allotted to each of these two phases of the experiment.

Results. The results of this experiment are summarized in Table I. As in the first experiment, the obtained cardiac *CR* was significant well beyond the 1% level of confidence at the end of acquisition. Although only 10.5 trials (median) were taken by this group to reach the avoidance-criterion (as compared with 16.0 in Experiment I)—possibly reflecting the acquisition of secondary reinforcing properties by the light—the difference between the two values is not significant. During the R_{AV} criterion-

TABLE I

MEAN CARDIAC RESPONSE (BEATS PER MIN.) DURING EACH OF THE THREE
PHASES OF EXPERIMENTS I AND II

(Last 11 basal determinations, last 5 conditioning trials, and
5 R_{AV} criterion trials)

Phase	Experiment I			Experiment II		
	Pre-tone	Post-tone	Diff.	Pre-tone	Post-tone	Diff.
Basal	80.8	79.4	1.4	77.7	77.9	−0.2
Conditioning	81.7	73.2	8.5*	78.1	70.5	7.6*
R_{AV} criterion	80.0	75.0	5.0*	77.5	75.8	1.7

* Significant beyond the 1% level of confidence by Wilcoxon's test for paired replicates. Individual data may be obtained from the American Documentation Institute.

trials, strength of cardiac *CR* had diminished to a non-significant level; moreover, 9 of the 10 *S*s (9/10 gives $P<0.01$) showed a smaller *CR* during R_{AV} *criterion*-trials than they did at the end of acquisition.[10] A possible explanation of this result is that the exteroceptive stimulus correlated with R_{AV} facilitates the development of a discrimination in which light comes to be the negative stimulus.

There exists the possibility, however, that the effect of the light may be unconditioned, and therefore not attributable to discrimination-learning. This possibility was explored during the third experiment.

---

[10] It may be noted that the *S*s of the two experiments did not differ significantly in cardiac response prior to avoidance-training. Neither, however, did they differ significantly in cardiac response during the avoidance-phase, which may be taken to imply that the introduction of the light failed to effect the cardiac *CR*. This conclusion might also be drawn from two other comparisons: first, with respect to the number of *S*s showing diminished *CR*s during R_{AV} criterion-trials as compared with the end of acquisition (the 15/20 in Experiment I does not differ significantly from the 9/10 in Experiment II); and secondly, with respect to the distributions of differences between the *CR*-magnitudes in the basal and R_{AV} criterion-stages which are not significantly different in the two experiments. Nevertheless, such an interpretation must be weighed against the fact that when the group in each experiment is taken as its own control, strength of *CR* during R_{AV} criterion-trials remains at a significant level in Experiment I but not in Experiment II.

EXPERIMENT III

The purpose of this experiment was to ascertain the unconditioned effect upon cardiac rate of the exteroceptive stimulus of Experiment II, and to determine whether that stimulus can assume discriminative control over cardiac *CR* through training by the method of contrasts.

Procedure. Eleven college students (men) served as Ss, each being taken, in a single session lasting about 2.5 hr., through three consecutive experimental phases: basal, acquisition, and discrimination-training. Before the start of the basal phase, and while the electrodes were being attached, each S was given the same initial instructions as in Experiments I and II. The basal phase consisted of 10 trials of tone-alone (T) and 10 trials of tone-light (TL), randomized separately for each S in such a manner that the separate unconditioned effect of T and TL upon heart-rate might later be evaluated. The tone lasted 1 sec., the light 1.5 sec.; on TL trials, the light was presented at the start of the fourth second after the tone. The acquisition-phase consisted of eight positive trials (T^+) in which tone was followed by shock, a procedure employed to produce a cardiac *CR* of some strength before the introduction of contrasting unreinforced trials associated with the negative stimulus. Since these trials followed basal trials without interruption, the first T^+ trial was considered a basal trial; similarly, the first discrimination-trial was considered another acquisition trial, 42 discrimination-trials being given in all. The discrimination-phase consisted of positive (T^+) and negative (TL^-) trials, duplicating the stimulus-shock correlations of Experiment II, with 7 T^+ trials randomly interspersed among the first 15 TL^- trials, 4 T^+ trials randomly interspersed among the last 15 TL^- trials, and with the last trial being positive (T^+) for all Ss. On all T^+ trials, the time-relations were the same as in the previous experiments (a 6-sec. shock administered 6 sec. after the end of the tone). On TL^- trials, the same light and arrangements were used as in Experiment II. The key used for R_{AV} in the two earlier experiments was not present for this experiment, since no R_{AV} was involved.

Results. The data indicate that the unconditioned effect of TL during basal trials was to decelerate heart-rate significantly. Thus, the pre-tone mean of the last 5 basal trials of 'tone-alone' was 75.2, the post-tone mean was 73.3, and the difference of 1.9 was, as usual in all previous work, non-significant. The pre-tone mean of the last 5 basal trials of tone-light was 75.2, the post-tone mean was 72.1, and the difference of 3.1 was significant beyond the 1% level of confidence.[11] This unconditioned cardiac-deceleration to TL is, of course, in the opposite direction to its effect upon cardiac *CR* in Experiment II, and strengthens the conclusion that the consequence of adding the light to R_{AV} in Experiment II was not an artifact but a genuine speeding of cardiac extinction.

The conclusion that the more rapid extinction in Experiment II resulted

---

[11] E. J. G. Pitman, *Notes on Non-Parametric Statistical Inference*, Department Mathematical Statistics, Columbia Univ., 1948.

from the discriminative action of the light receives further confirmation from the data of Fig. 1 which shows the cardiac response to T^+ and TL^- during the discrimination-training. The values plotted are averages for every two sequential T^+ trials and for every block of five sequential TL^- trials, reflecting the 2:5 ratio of frequencies of the two presentations. The
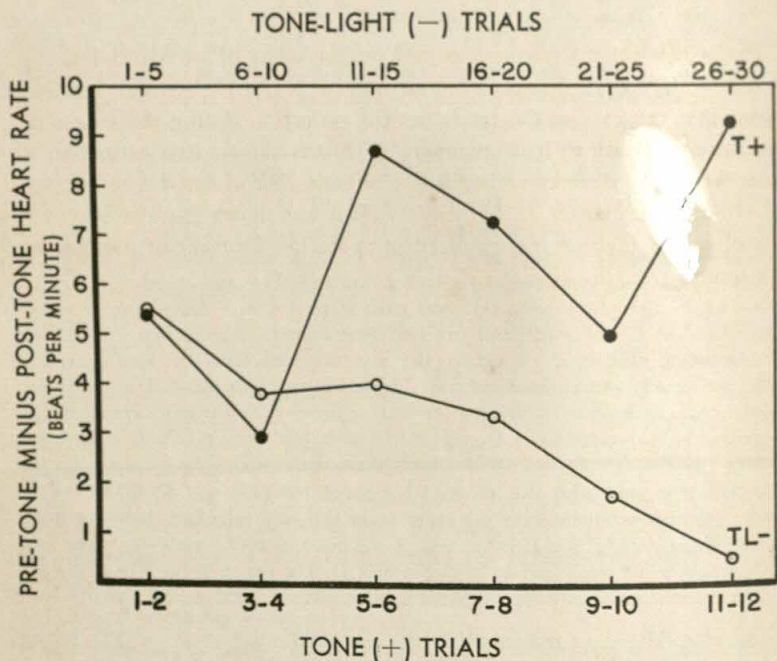
FIG. 1. DISCRIMINATIVE CONTROL OVER A CARDIAC CR ESTABLISHED BY THE METHOD OF CONTRASTS IN EXPERIMENT III.
T^+ values are averages for two sequential trials; TL^- values are averages for five sequential trials.

cardiac CR seems to generalize well from T^+ to TL^- at the end of acquisition and before differential reinforcement. The cardiac drop to the negative stimulus during its first five presentations during the discrimination-phase is significantly greater than during the last five trials of the basal phase ($0.02 < P < 0.05$), thus indicating a genuine generalization effect. A typical progressive separation soon occurs, however, in which the positive stimulus produces highly significant responses while CR to the negative stimulus diminishes to a level statistically indistinguishable from zero. Data not included in the table indicate, in fact, that the negative stimulus

eventually tends to produce cardiac acceleration. There seems little doubt, in view of these findings, that discriminative control by an exteroceptive stimulus over the cardiac CR is possible, and, consequently, that the light accompanying R_{AV} in Experiment II might have, as argued there, served this added discriminative function.

SUMMARY

Three experiments were performed on the effect of avoidance-training on a conditioned cardiac-response in human Ss. In Experiment I it was shown that the cardiac CR tends toward extinction during the course of avoidance-responding. In Experiment II it was shown that extinction is faster when an exteroceptive signalling stimulus is correlated with successful avoidance-responses. In Experiment III it was shown that the effectiveness of such signalling may be ascribed to the development of discriminative control over the cardiac CR.

FACTORS AFFECTING ESTIMATION OF DEPTH WITH VARIATIONS OF THE STEREO-KINETIC EFFECT

By GLORIA J. FISCHER, Purdue University

A depth-effect can be produced with certain plane figures when they are attached to a rotating disk. Apparent movement of the units also is perceived. According to Musatti,[1] the first demonstration of a stereo-kinetic effect from a plane figure was made by Benussi, using two circles. More recently, Wallach has demonstrated the effect using four circles.[2] The figure was similar to *b* in Fig. 1 and seen as a cone, when rotated at appropriate speeds. The kinetic effect is apparent rotation of the figure.

In preliminary studies, variations such as 'a' through 'e' in Fig. 1 were studied. The apparatus consisted of a motor-driven disk, 22 in. in diameter, to which the stimulus-figures were attached. For all the figures the speed of rotation was 60 r.p.m. Significantly more than half the *O*s tested perceived figures 'b' and 'd' in Fig. 1 as having depth, *i.e.* as a cone and a cylinder, respectively, whereas figures 'a' and 'c' were perceived as flat. Some *O*s perceived figure 'e' as having depth. When so perceived, this figure had the striking appearance of a cylinder, twisting and snake-like.

These studies raised the following questions: (*1*) What factors determine whether depth is perceived? (*2*) What factors or parameters determine the amount of depth perceived? (*3*) What individual differences exist with respect to this visual experience? The second question received primary emphasis in the present study, although certain aspects of the design provided information relating to all three questions.

METHOD. A list of the factors considered in the experiment are included in Table I. Off-set refers to the spatial separation of the two circles. Since neither concentric nor tangential circles were found to produce depth, it seemed possible that these conditions were limits, within which off-set might be a determinant of depth. Circles of both equal and unequal size were used. The resulting stimulus-figures, drawn to scale, are shown in the second row of Fig. 1. It will be noted that for each level of off-set, the absolute amount of separation was greater for the cone (f through h) than cylinder (i through k), since, for the cone, off-set was made relative to the radius left of the larger circle (*r*). Position refers to the locus of the figure on the disk. In 'central' position the figure was attached to the middle of the disk. In the 'peripheral' positions the figure was centered on a radius of the disk. For unequal rates of motion, the plane of off-set was placed along the horizontal radius of the disk, and for equal rates the plane of off-set was per-

[1] C. L. Musatti, *Sui fenomeni stereocinetici, Arch. Ital. Psicol.,* 3, 1924, 105-120.
[2] Personal communication.

pendicular to the horizontal radius of the disks. A comparison of peripheral levels comprised a test of the saliency of 'movement perspective'[3] as a possible cue to the perception of depth. A comparison between cones and cylinders similarly tested the saliency of a size-difference cue. The additional factors, manner of viewing and direction of motion of the disk are self-explanatory. To study individual differences, the experiment was repeated with Os having poor stereoscopic vision. These Os,
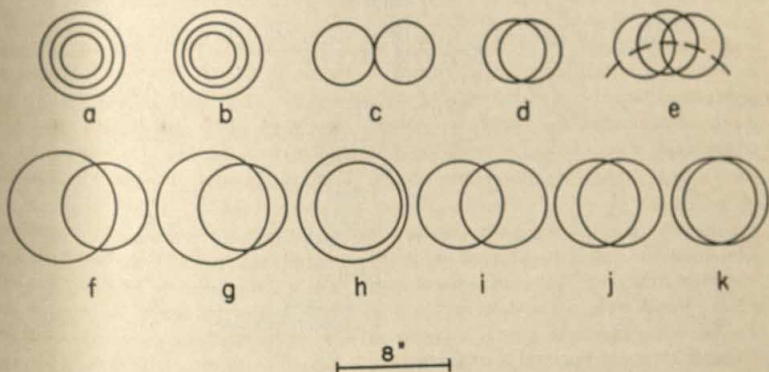
FIG. 1. STIMULUS-FIGURES.
(a–e, used in preliminary studies; f–k, used in the main experiment)

which will be referred to as abnormal, were selected from the individuals with 20/200 corrected vision in one eye and essentially normal vision in the other.

The stimulus-figures were revolved in the manner indicated. O estimated the amount of depth perceived by adjusting a sliding gauge from a zero-point along a scale which was visible only to E. The estimate was recorded by E to the nearest 1/16. Verbal observations also were recorded, as to which of the two circles in a given treatment combination appeared nearer to O.

The statistical design was a $2^3 \times 3^2$ mixed factorial with confounding. A third-order interaction (1 df.) was confounded with blocks of 36 observations each.[4] An O was used in the experiment only if he perceived depth in the first stimulus-figure presented to him. Eight undergraduates with no obvious visual defects were selected in this manner to provide four replications of the experiment. Two abnormal Os allowed a fifth replication of the experiment. For monocular viewing all Os were permitted to use the preferred eye. The analysis on the measured variable was carried out for three groups: the normal group, the abnormal group, and the total group.

Results. The three groups were found to differ with respect to only one factor. Consequently, for brevity, data on the measured variable are given for the normal group alone, and group-differences are described in the text. Table I shows the mean amount of perceived depth, in inches, in relation to all main variables. The

[3] R. S. Woodworth, *Experimental Psychology*, 1953, 652.
[4] Oscar Kempthorn, *The Design and Analysis of Experiments*, 1953, 390-426, esp. 419.

influence of off-set and viewing were statistically significant beyond the 5% level of confidence. (Off-set was a significant effect for both the normal and abnormal groups.) As can be seen from the Table, the relationship between proportionate off-set and amount of perceived depth is a positive one. A test for trend, using the Duncan process and Tables,[5] indicated further that deviations from linearity were insignificant.

TABLE I

MEAN AMOUNT OF PERCEIVED DEPTH UNDER VARIOUS CONDITIONS FOR THE NORMAL GROUP

Off-set*	Am't depth	St'd. error	Viewing*	Am't depth	St'd. error
r/3	4.52	± .13	monocular	6.89	± .19
2r/3	6.46		binocular	5.61	
r	7.76				
			Figure		
Placement			cone	6.56	± .19
central	6.09	± .17	cylinder	5.93	
peripheral motion equal	6.31				
			Motion		
peripheral motion unequal	6.34		clockwise	6.14	± .15
			counterclockwise	6.32	

* Significant at the 5% level of confidence.

Viewing was a significant effect for the normal group only. Table 1 indicates that monocular viewing produced the greater appearance of depth. One main effect, 'figure,' was significant beyond the 5% level for all Os combined, but not in either group alone. Two interaction-effects were significant beyond the 5% level: off-set × figure in the normal group only, and off-set × viewing × groups. Relevant individual contrasts were made, again using the Duncan process and Tables. Results for the off-set × figure interaction indicated that the cone tends to be perceived as having more depth than the cylinder, and to about the same extent when the off-set is r/3 or 2r/3, but that the difference does not appear when the off-set is as great as r. Results for the off-set × viewing × groups interaction indicated that the difference in perceived depth between the cone and cylinder decreases for normal Os as they go from monocular to binocular viewing, while the perceived difference increases for abnormal Os as they go from monocular to binocular viewing.

Verbal reports from all Os indicating which of the two circles was perceived as nearer may be summarized briefly. It should be noted that Os reported many reversals with the cone and some with the cylinder, and in these instances O was asked which circle was more often seen as nearer during a given treatment-combination. For the cone, the smaller circle was nearer in 68% of the cases. For the cylinder, the outer circle was nearer in 92% of the cases with unequal motion, and the leading circle was nearer in 68% of the cases with equal rate of motion.

Discussion. The findings will be discussed, first as they relate to possible determinants of the depth-effect, and secondly as they relate to possible determinants of the amount of perceived depth. With respect to possible determinants of the

[5] D. B. Duncan, Multiple range and multiple F tests, *Biometrics,* 2, 1955, 1-42.

effect, a differential-size cue can clearly be rejected as salient. Preliminary findings indicate perceived flatness where differential size is the only cue present (as with concentric circles.) Moreover, the depth-effect does appear where such a cue is not present (where circles of equal size are used and a cylinder is seen.) Similarly, the findings suggest that movement-perspective also can be rejected. When differential rate of motion of the two circles is the only cue present (as in tangential circles), the effect does not appear. Yet the effect does appear where such a cue is not present. Movement-perspective appears to be a relevant cue, however, since the circle moving at the greater rate is seen as nearer to O a large percentage of the time.

A probable determinant of depth would appear to be off-set or simple spatial separation of the units in the figure, since depth and kinetic effects are produced *only* under such conditions. A possible determinant which cannot be rejected in the present study is perspective. Such a cue might be considered to be automatically engendered by off-set, due to the tapering effects at points of contact of the circles. In this sense perspective and off-set cues cannot be separately evaluated here. A test of whether the factor really is perspective would be to remove the tapering effects by moving congruent off-set squares back and forth along a plane. If depth were perceived, the necessity of a perspective-cue could be rejected. Such an outcome is rendered probable, perhaps, by the fact that such off-set squares can be seen as a cube without rotation. Tentatively, then, the hypothesis is proposed that off-set and physical rotation comprise the physical factors necessary to produce a stereo-kinetic effect. Given these circumstances, the inclusion of relevant cues, such as movement perspective, may facilitate the effect, *i.e.* produce it in more Os.

With respect to possible factors affecting the *amount* of perceived depth, both off-set and manner of viewing were found to be significant. In the present experiment, amount of perceived depth increased with off-set. The greater perceived depth for the cone than for the cylinder can be accounted for by the greater off-set in the former case. That the latter finding appeared for the total group only is, perhaps, clarified by consideration of the off-set $\times$ figure interaction. The cone tended to be seen with greater depth than the cylinder, and to about the same extent, by the normal group when the off-set was $r/3$ or $2r/3$, but the difference did not appear when the off-set was as great as r. Thus r off-set of the cone, (4 in.—the size of the circles used in the present study) appears to have exceeded the absolute amount of depth possible of resolution in the given figures. Imposing an absolute limit on r for a given figure then, the general relationship between off-set and perceived depth, through an amount of off-set equal to r, would appear to be adequately described by a linear function. In this respect it should be noted that with figures containing more than two circles (*e.g.* 'e' in Fig. 1) off-set would correspond to the total spatial separation from first to last circle, and the linear relation to perceived depth would be expected to hold as long as no single off-set from one circle to the next exceeded an r amount.

The remaining factor found to affect the amount of perceived depth was the manner of viewing. Monocular viewing was found to produce greater perceived depth than binocular viewing, but in the normal group only. A possible explanation might be that a lack of disparity with binocular viewing may function as a

cue of 'flatness.'[6] Such a cue might be expected to reduce apparent depth. For the abnormal Os, who evidenced poor stereoscopic vision, the flatness-cue may be less functional, due to their relative inexperience with it, as compared to the Os with normal stereoscopic vision.

A related finding, the off-set $\times$ viewing $\times$ groups interaction, is difficult to interpret and presents an inconsistency with the proposed explanation of the major finding with regard to viewing. It will be remembered that the difference between the amount of depth perceived in the figure with the greater amount of off-set (cone) and the one with the lesser amount of off-set (cylinder) decreased in the normal Os from monocular to binocular viewing, as it *increased* in the abnormal group from monocular to binocular viewing. The expectation for the abnormal group would be, on the contrary, that the difference would remain relatively constant or, perhaps, decrease slightly.

Although off-set and manner of viewing are clearly determinants of the amount of perceived depth, it should be noted that off-set *per se* can not be a sufficient determinant. Any figure in the present study could be perceived in depth at an infinite number of possible angles of inclination from the frontal plane, and the corresponding amount of perceived depth would vary. Thus it is apparent that neither the amount of perceived depth nor the stereo-kinetic effect is entirely explicable in terms of physical conditions alone. A possible explanation is that, given the necessary physical determinants, *i.e.* rotation of off-set units, individuals must organize the units into three-dimensional objects in a manner consistent with their experience. That is, apparent depth and apparent movement of the circles (units of the present study) at the rate of one revolution per physical revolution of the disk may be effects determined by the suggested physical circumstances of the stimulus-situation. Perception of revolving figures, on the other hand, is an organization in accordance with the O's experience.

Consistent with experience, the physically determined effects would be apparent in a real object, such as a cone or cylinder, if it were being seen as essentially rotating about its base or actually revolving about a point on its axis. Only in the latter case would the top and base appear to describe symmetrical circles as the object revolved. For the apparent revolution of the circles to appear as such object-revolution, the angle at which the object would be perceived, consistent with experience, would have to be an angle sufficiently central from the frontal plane. Thus, this experiential factor plus a given amount of off-set would be sufficient to determine the amount of perceived depth! Although a cone might need to be perceived at a less central angle than a cylinder, the angle would be constant for each type of figure. Increased off-set (perceived, of course, as the side of the figure) would result in increased apparent depth. Further, for circles or units of a given size this explanation would predict maximum perceived depth at an r amount of off-set from one unit to the next, and so forth, since increasing off-set beyond r would require the object to be perceived at progressively less central angles, in accordance with the restriction that progressively less or, most important, less than *all* of the base would be visible upon rotation. With such a restriction the stereo-kinetic effect should not even be possible, since apparent movement would not appear as object-rotation at the less central angles.

[6] Harold Schlosberg, Stereoscopic depth from single pictures, this JOURNAL, 54, 1941, 601.

In conclusion then, within suggested limits, off-set or spatial separation of visual units and their physical rotation are proposed as factors necessary to produce the stereo-kinetic effect. It is suggested, however, that O must be able to organize apparent movement of the units into a three-dimensional object consistent with his experience. Where the effect is present, off-set is linearly related to the amount of perceived depth, and depth is enhanced by monocular viewing. If valid, the experiential factor proposed would render off-set a sufficient determinant of the amount of perceived depth. With respect to the finding on differential viewing, a proposed explanation is based on the presence or absence of a binocular cue of flatness.

SUMMARY

A stereo-kinetic effect was demonstrated by revolution of off-set circles of both equal and unequal size. The effect was realized through off-set or spatial separation of two circles as great as the radius of the largest circle in a given figure. Factors found to affect the amount of perceived depth were amount of off-set and manner of viewing. Apparent depth was linearly related to the amount of off-set used, and monocular viewing enhanced perceived depth.

The findings were evaluated with respect to factors determining the perception of depth and the amount of perceived depth. An explanation of the stereo-kinetic effect in terms of both physical and experiential factors was proposed.

THE EFFECT OF TARGET–VELOCITY UPON THE TRIGONOMETRIC RELATIONSHIP OF PRECISION AND ANGLE OF LINEAR PURSUIT-MOVEMENTS

By Richard F. Thompson, James F. Voss, and W. J. Brogden, University of Wisconsin

This paper presents the results of an experiment on the effect of target-velocity upon the relationship of precision and angle of linear pursuit-movements. Previous studies have found the latter relation to be described by the following trigonometric equation :[1] $y = a + b \cos 2x + c \sin 2x$, in which $y =$ precision of right-arm movements in terms of group mean frequency of stylus-contacts, $x =$ angle from the body at which the movement is made,[2] $a =$ the constant that determines the baseline of the curve (mean frequency of stylus-contacts for all angles), and b and c are the constants in terms of which the amplitude (d) and the phase-angle ($2e$) of the curve are determined $[d = (b^2 + c^2)^{1/2}; \cos 2e = c/d]$. Earlier studies also have shown that the trigonometric relationship of precision and angle of linear pursuit-movements is maintained over the ranges studied for other variables such as practice, bilateral transfer, laterality, angle of tilt, and angle of slant. In all cases these variables have produced systematic quantitative changes in some of the terms of the equation and thus provided information about secondary parametric relations within the primary trigonometric relationship of precision and angle of linear pursuit-movements. Target-velocity has not heretofore been varied, but has been held constant throughout at 3.0 cm./sec. In light of the results from other experiments in this series, variation in target-velocity might be expected to alter systematically some of the terms of the trigonometric equation, but not to destroy the basic trigonometric relationship of precision and angle of linear pursuit-movements. Since studies of other kinds of tracking behavior in which target-velocity has been the independent variable have uniformly found precision in tracking to decrease

[1] R. E. Corrigan and W. J. Brogden, The effect of angle upon precision of linear pursuit-movements, this JOURNAL, 61, 1948, 502-510; The trigonometric relationship of precision and angle of linear pursuit-movements, *ibid.*, 62, 1949, 90-98; Brogden, The trigonometric relationship of precision and angle of linear pursuit-movements, this JOURNAL, 66, 1953, 45-56; G. E. Briggs and W. J. Brogden, Bilateral aspects of the trigonometric relationship of precision and angle of linear pursuit-movements, this JOURNAL, 66, 1953, 472-478; G. E. Briggs, R. F. Thompson, and W. J. Brogden, The effect of angle of tilt upon the trigonometric relationship of precision and angle of linear pursuit-movements, this JOURNAL, 67, 1954, 475-483; Thompson, and Brogden, The effect of angle of slant upon the trigonometric relationship of precision and angle of linear pursuit-movements, this JOURNAL, 68, 1955, 615-623.

[2] Angle is designated by Cartesian coördinates: 0° represents the position in which the track is normal to the frontal plane of the body of S, and the arm-movement is started with the stylus close to, and continued away from, the frontal plane of the body.

as target-velocity is increased,[3] a similar effect of target-velocity on linear pursuit-movements might be expected here. The hypothesis under test in the present experiment, therefore, is that variation in target-velocity will change the over-all precision of linear pursuit-movements (*a*, the baseline) but will not otherwise affect the trigonometric relationship.

Method. The basic unit of the experimental design is an 8 × 8 latin square in which angle from the body is represented by the latin letters, sequence of angles is represented by the rows of the square, and ordinal position of angles is represented by the columns. The 10 replications of this square are represented by 5 target-velocities (2.5, 3.0, 3.5, 4.0, and 4.5 cm./sec.) and by 2 *Es*. In other words, each of the 2 *Es* completed 5 replications of the square, 1 for each of the 5 velocities. Each of the 80 *Ss* tracked at each of the 8 angles, but at only 1 velocity. The angles are 0°, 30°, 45°, 60°, 90°, 120°, 135°, and 150°.

The apparatus is the same as that described previously.[4] It consists of two brass plates, 35.0 cm. in length and 4 cm. apart, that rest upon a glass plate. A small cylindrical target moves at constant velocity under the glass plate, and *S* matches the target movement by movement of a stylus on the surface of the glass plate. Errors are counted whenever *S* hits either of the brass plates. Contact with a plate closes the input-circuit of a Potter Electronic Counter, Model 67. Target-velocity is manipulated by varying the frequency-output of the oscillator that provides power for the synchronous motor that drives the target by a cord-and-pulley arrangement. The frequencies of the oscillator are 50, 60, 70, 80, and 90 c.p.s. respectively for the target-velocities of 2.5, 3.0, 3.5, 4.0, and 4.5 cm./sec. The extreme velocities of 2.5 and 4.5 cm./sec. are just within the limitations of the apparatus.

The instructions to the *Ss* and the procedure were the same as those used in previous studies in this series.[5] The *Ss* were University of Wisconsin students (men) enrolled in the course in elementary psychology.

Results. This analysis is the extension of the usual analysis of variance of the latin square, previously used for similar experimental designs in this series of studies, in which not only are angle, ordinal position of angle, and sequence of angle, orthogonal variables, but experimenter and velocity of the target also are orthogonal, both to one another and to the other three variables.[6] Tests of significance were made separately for uncorrelated data, between-*Ss* variation, and correlated data, within-*Ss* variation. Of the total sum of squares, 59.0937, with 639 *df*, a between-*Ss* sum of squares of 21.0943 with 79 *df*. and a within-*Ss* sum of squares of 37.9974 with 560 *df*. were calculated. Further fractionation of each of these two sources is represented in Table I by the mean square for

[3] D. G. Ellson, H. Hill, and F. Gray, Wave length and amplitude characteristics, Aero Medical Laboratory, U. S. Air Force, Headquarters, Air Material Command, Engineering Division, Serial No. TSEAA-694-2D, 1947, 1-23; J. F. Voss, The effect of target brightness and target speed upon tracking proficiency, *J. exp. Psychol.*, 49, 1955, 237-243.

[4] Corrigan and Brogden, *opp. citt.*, 61, 1948, 102 f.; 62, 1949, 90 f.

[5] Corrigan and Brogden, *op. cit.*, 61, 1948, 102 f.

[6] Briggs, Thompson, and Brogden, 475 f.; Thompson and Brogden, *op. cit.*, 615-623.

each fractional source. The $E \times V \times SA$ interaction is the basic error-term against which all fractional mean squares appearing above it in the table were tested. The only F-ratio of statistical significance at the 5% level of confidence is that for E. The mean squares for fractional sources for the within-Ss sum of squares were all tested against the $RS \times E \times V$ interaction-term. Statistically significant F-ratios at the 5% level of confidence were obtained for angle, $A \times E$, ordinal position, $OP \times E \times V$, and $RS \times V$. In terms of the variables manipulated ex-

TABLE I

Results from Analyses of Variance

Source	df	Standard error-score		Adjusted error-score		F at 5% level
		Mean square	F	Mean square	F	
Total between-Ss	79	—	—	—	—	—
Experimenter (E)	1	4.2736	22.14	0.0401	19.10	4.20
Velocity (V)	4	0.1186	0.61	0.2436	116.00	2.71
Sequence of angle (SA)	7	0.2960	1.53	0.0031	1.48	2.36
E×V	4	0.0755	0.39	0.0006	0.29	2.71
E×SA	7	0.2942	1.52	0.0028	1.33	2.36
V×SA	28	0.2325	1.20	0.0021	1.00	1.91
E×V×SA	28	0.1930	—	0.0021	—	—
Total within-Ss	560	—	—	—	—	—
Angle (A)	7	2.6192	102.31	0.0254	84.67	2.09
A×E	7	0.1442	5.63	0.0013	4.33	2.09
A×V	28	0.0296	1.16	0.0003	1.00	1.61
A×E×V	28	0.0240	0.94	0.0002	0.67	1.61
Ordinal position (OP)	7	0.2322	9.07	0.0021	7.00	2.09
OP×E	7	0.0262	1.02	0.0003	1.00	2.09
OP×V	28	0.0275	1.07	0.0002	0.67	1.61
OP×E×V	28	0.0437	1.71	0.0004	1.33	1.61
Residual for square (RS)	42	0.0268	1.05	0.0002	0.67	1.49
RS×E	42	0.0254	0.99	0.0002	0.67	1.49
RS×V	168	0.0409	1.60	0.0003	1.00	1.35
RS×E×V	168	0.0256	—	0.0003	—	—

perimentally, the significant interaction of $OP \times E \times V$ is obscure and that of $RS \times V$ is not meaningful. The significance of angle and ordinal position follow the results of previous experiments and represent respectively the trigonometric relationship of precision and angle of linear pursuit-movements, and a practice effect from the first to the last of the eight angles at which pursuit-movements are measured serially, represented by the successive columns of the latin square.

The significant variations due to E and the interaction of A and E have not been found before. Plots of the data for each E, and of the trigonometric regression equation, $y = a + b \cos 2x + c \sin 2x$, fitted to the data for each E by the method of least squares, are presented in Fig. 1. Examination of the figure shows that the Ss tested by the two Es differ in over-all precision of linear pursuit-movements (represented by the significant E-term) and that, for the poorer group of Ss, performancce is relatively worse at the most difficult angles than it is for the other group of Ss (represented by the significant A-by-E interaction). Since no reasonable hypothesis could be formulated to account for these differences as a function of a

difference between Es, the possibility of sampling variation was considered. The regression equation for E_1 is $y = 1.91 + 0.16 \cos 2x - 0.06 \sin 2x$, and for E_2 is $y = 2.08 + 0.24 \cos 2x - 0.11 \sin 2x$. The phase-angle, baseline, and amplitude terms are respectively 20° 55', 1.91, and 0.15 for E_1 and 24° 48', 2.08, and 0.21 for E_2. Examination of the terms of the trigonometric regression equations fitted to data obtained in previous studies in this series shows the values for the terms of the equations for the two Es of the present experiment to fall inside the range
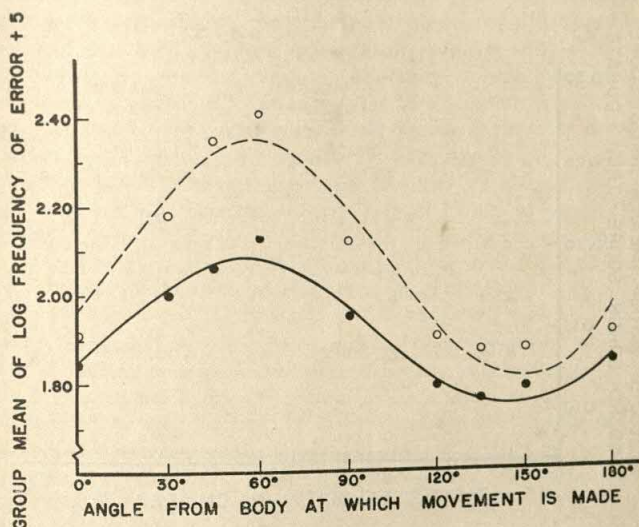
FIG. 1. THE TRIGONOMETRIC RELATIONSHIP OF PRECISION AND ANGLE OF LINEAR PURSUIT-MOVEMENTS

The closed circles are the means for Ss tested by E_1 and the open circles are the means for Ss tested by E_2. Each curve is the least-squares fit to the empirical data of the trigonometric equation, $y = a + b \cos 2x + c \sin 2x$.

for the appropriate term in every case. It appears reasonable, therefore, to attribute the obtained differences betwen Es to sampling variations of the Ss.

The failure of analysis of variance to show any significant effect for velocity may be due to the confounding of this variable with that of trial-duration. As velocity is increased, the time spent on each trial decreases. Trial-duration is 14.0, 11.7, 10.0, 8.8, and 7.8 sec. respectively for the velocities of 2.5, 3.0, 3.5, 4.0, and 4.5 cm./sec. Since analysis of total errors per trial shows no variation as a function of velocity, it then follows that errors per unit time must increase as velocity increases. To equate the measures of error per unit time at the different velocities, the error-score for each trial for every S was divided by the duration of the trial. This adjustment of the error-score makes 1 sec. the unit of time for each of the velocity conditions and provides for a comparison of precision of linear pursuit-movements at the different velocities on a uniform scale. The results of analysis of variance of the adjusted error-score are given in the columns of Table I headed *Adjusted error-score*. The results are similar to those obtained from analysis of

the standard error-score except that velocity is now found to be a significant source of variation, and the interactions involving velocity ($OP \times E \times V$ and $RS \times V$), that were statistically significant in the previous analysis, are not significant. The discussion of the results of analysis of the standard error-score holds for the results of analysis of the adjusted error-score, particularly with regard to E as a source of variation. In dealing with the relation between velocity and adjusted error-score, however, it is desirable to separate the results for the Ss tested by each of the two Es.
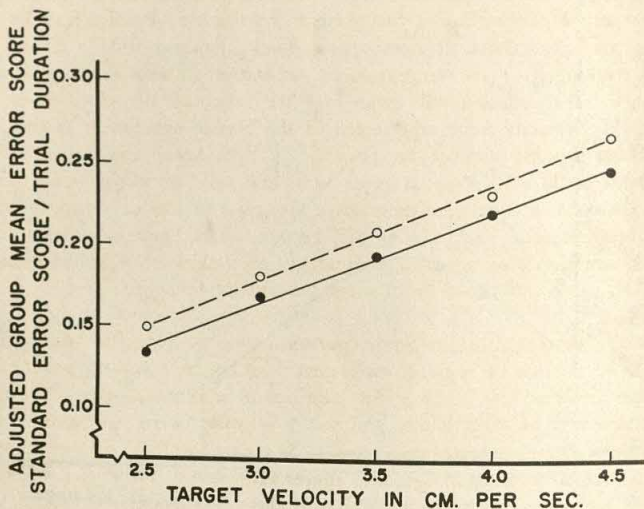
FIG. 2. THE RELATION BETWEEN PRECISION OF LINEAR PURSUIT-MOVEMENT AND TARGET-VELOCITY

The closed circles are the means for Ss tested by E_1 and the open circles are the means for Ss tested by E_2. The lines represent the least-square regression-equation of the straight-line function, $y = ax + b$.

Group mean adjusted error-score computed over all angles is plotted in Fig. 2 as a function of velocity for each E. Since the relation between adjusted error-score and velocity is linear in both cases, a linear regression equation was fitted to each set of data by the method of least squares. These equations are $y = 0.0542x + 0.0015$ for E_1 and $y = 0.0560x - 0.0110$ for E_2. Both equations are plotted in Fig. 2.

The trigonometric relationship between precision and angle of linear pursuit-movement holds for the adjusted error-score in much the same way as it does for the standard error-score. Since there is no significant interaction of angle and velocity, the form of the trigonometric relation does not differ as a function of velocity. The baseline (mean adjusted error-score over all angles) does, of course, differ as a function of velocity, showing the linear relationship presented in Fig. 2.

Discussion. The hypothesis examined in the present experiment was, in the main, confirmed. Target-velocity has a significant effect upon precision of linear pursuit-movements, but only when measurement of error is adjusted for the different velocities by dividing total error per trial by the trial-duration. Error-frequency

per unit of time increases as a linear function of velocity. Velocity is subsidiary to angle from the body, in that the trigonometric relationship between precision and angle of linear pursuit-movements is maintained for all the velocities used in the experiment. Velocity may be represented as a parameter of the trigonometric regression equation since the baseline, a (mean frequency of stylus-contacts for all angles adjusted for trial-duration), increases as a linear function of velocity.

The effect of velocity on tracking-error found in this experiment is similar to that obtained in other studies.[7] It is impossible to make any direct comparisons between these experiments and the present one because of differences in method, variables, and parameters. It does appear likely, however, that a linear relation between tracking-error and target-velocity holds over a wide range of conditions. Whether a linear relation will hold over the range of all velocities is another question. The velocity-range represented in the present experiment is considerable, but tracking would certainly be possible at both lower and higher velocities. Inability to track will occur at some very low velocity and at some very high velocity. It is likely that near these extremes, prior to loss of ability to track, the relation between error and velocity will be non-linear. These hypotheses are based on the assumption that adequate comparison of tracking-error at different velocities can be made only in terms of error per unit of time.

Summary. Performance of linear pursuit-movement with the right arm was studied as a function of angle of movement from the body and target-velocity. The experiment involved use of an 8×8 latin square repeated 10 times. The 8 angles were represented by latin letters and the 5 velocities were represented by the 2 squares. Each of 2 Es collected data from 40 Ss assigned to the 5 squares representing velocity. Thus, E is orthogonal to both angles and velocity. The angles were $0°$, $30°$, $45°$, $60°$, $90°$, $120°$, $135°$, and $150°$, and the velocities were 2.5, 3.0, 3.5, 4.0, and 4.5 cm./sec.

Analysis of variance of the standard error-scores revealed the trigonometric relationship of precision and angle of linear pursuit-movements as found in previous studies in this series. There was a significant difference in the functions obtained from the data of the two Es that appears to be due to sampling variation of the Ss. Velocity has no significant effect, but when the error-scores are divided by trial-duration, thus reducing the scores to errors per unit time, this measure increases significantly as velocity increases. The function is linear over the range of velocities used in the experiment.

[7] Ellson, Hill, and Gray, *op. cit.,* 1-23. Voss, *op. cit.,* 237-243.

THE EFFECT OF PARTIALLY DELAYED REINFORCEMENT AND TRIAL-DISTRIBUTION ON THE EXTINCTION OF AN INSTRUMENTAL RESPONSE

By Earl D. Scott and Edward L. Wike, University of Kansas

In 1951, Crum, Brown, and Bitterman reported an investigation which has not received sufficient attenion.[1] Rats were trained to traverse a runway, one group being rewarded immediately on every trial, while a second group was rewarded immediately on half the trials and after a delay of 30 sec. on the remaining half (partial delay of reinforcement). In subsequent tests, the latter group was found to extinguish less rapidly than the former; that is, partial delay of reinforcement increased resistance to extinction. This outcome appeared to contradict the Hull-Sheffield interpretation of partial reinforcement in terms of the carryover the after-effects of reinforcement from trial to trial.[2] Since both groups of rats had been rewarded on every trial, it was argued, the difference in rate of extinction must have been due to events taking place *within* the training trials.[3]

The primary purpose of the present experiment was to replicate the finding of Crum, Brown, and Bitterman, which has the same far-reaching implications for behavior-theory as the paradox of partial reinforcement itself. A second purpose was to study the effect of distribution of practice, a variable which has led to inconsistent results in work on partial reinforcement.[4]

Subjects. Fifty-nine naïve hooded female rats of the Long-Evans strain were studied. They were bred in the laboratories of the departments of psychology and physiology. Eleven of the animals were discarded in the early stages of the experiment, six which displayed excessive emotionality, and five others (randomly selected) to equalize the number of Ss in each treatment. The animals were assigned at random to each of four training groups of 12 Ss (M-100, M-50, S-100 and S-50),

---

[1] Janet Crum, W. L. Brown, and M. E. Bitterman, The effect of partial and delayed reinforcement on resistance to extinction, this Journal, 64, 1951, 228-237.

[2] V. F. Sheffield, Extinction as a function of partial reinforcement and distribution of practice, *J. exp. Psychol.,* 39, 1949, 511-549; C. L. Hull, *A Behavior System,* 1952, 120ff.

[3] Other evidence against the carry-over theory is provided by experiments on alternating reinforcement. *Cf.* E. D. Longenecker, John Krauskopf, and M. E. Bitterman, Extinction following alternating and random partial reinforcement, this Journal, 65, 1952, 580-587; D. W. Tyler, E. C. Wortz, and M. E. Bitterman, The effect of random and alternating partial reinforcement on resistance to extinction in the rat, this Journal, 66, 1953, 57-65.

[4] *Cf.* Sheffield, *op. cit.,* 511-549; D. A. Grant, J. P. Hornseth, and H. W. Hake, The influence of the inter-trial interval on the Humphreys' 'random reinforcement' effect during the extinction of a verbal response, *J. exp. Psychol.,* 40, 1950, 609-612; and S. Weinstock, Resistance to extinction of a running response following partial reinforcement under widely spaced trials, *J. comp. physiol. Psychol.,* 47, 1954, 318-322.

each of which was further divided into two subgroups, making a total of eight treatments with six Ss per treatment.

Apparatus. The elevated runway was a duplicate of the one used by Crum, Brown, and Bitterman.[5] It consisted of a narrow runway 9 ft. long, a delay-box, and a feeding compartment. The only difference between the original apparatus and the one employed in the present study was that original runway was 20 ft. long.

Procedure. A careful attempt was made to duplicate the procedure reported by Crum, Brown, and Bitterman. Throughout the study, the animals were kept on a 24-hr. feeding schedule, which was initiated a week prior to the beginning of the preliminary training trials. During this period, the animals were handled daily and afterwards were fed wet Purina mash for 1 hr. in their home cages. After several days of being fed small amounts of food in the goal-box, each animal was given five massed and immediately reinforced preliminary training trials on the runway, and five additional trials of the same kind were given on the following day. The reinforcement throughout the experiment consisted of access to the food tray for 10 sec.

On the next 10 days, each animal was given 10 training trials per day. Groups M-100 and S-100 were immediately reinforced on all trials. Groups M-50 and S-50 were partially delayed, *i.e.* rewarded immediately on five of the daily trials but were forced to wait for 30 seć. before being rewarded on the other five trials. The occurrence of immediate and delayed reinforcement followed quasi-random patterns with two restrictions: (1) the first and last daily trial were always immediately reinforced; and (2) no more than two immediately rewarded or two delayed trials occurred in succession. Groups M-100 and M-50 received massed training. The interval between trials was 15 sec., during which time S was confined in a waiting cage. To maintain a constant hunger drive and equalize the time between trials, the animals of these two groups were run in pairs. The training of Groups S-100 and S-50 was spaced. Their treatment was identical to that of the massed groups, except that the interval between trials was never less than 15 min., and the animals were run alternately and in rotation.

On each of the two days after the training series, half of the animals of each of the four training groups were given 10 extinction-trials. On these trials, each animal was confined in the delay-chamber for 10 sec., and if an animal failed to enter the chamber within a period of 3 min. it was placed in the delay-box for the customary 10-sec. period. Half of the Ss in each training group were extinguished under massed conditions and half under spaced. The same inter-trial intervals, 15 sec. and 15 min., were utilized during extinction.

Performance in all stages of the experiment was measured in terms of running-time. A stopwatch was started when the rat was placed on the runway and stopped as it entered the delay-box.

Results. To determine whether the four training groups were initially equivalent, the running-times for the 10 preliminary trials, in which all Ss were run under identical conditions, were transformed to common logarithms and a simple analysis of variance was performed. The resulting F-ratio of 1.025, with 3 and 44 df. was not significant.

---

[5] Crum, Brown, and Bitterman, *op. cit.,* 233.

The course of acquisition for the four training groups is shown in Fig. 1. The four groups were compared at the termination of training by transforming the
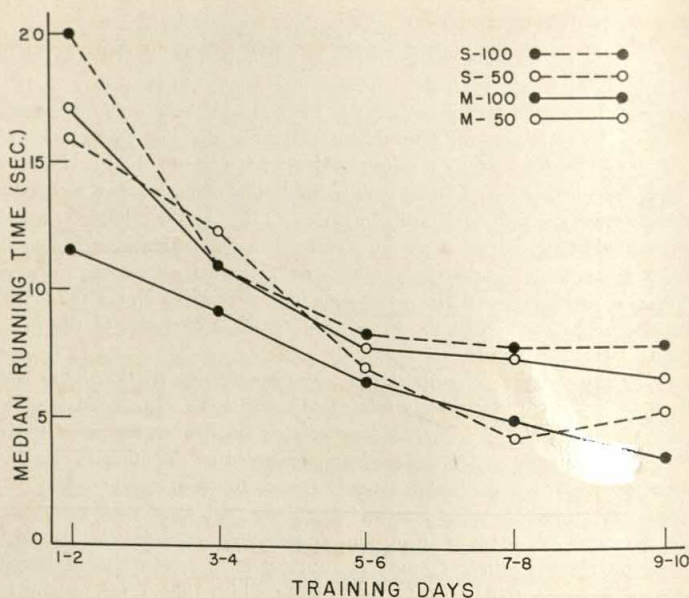
FIG. 1. ACQUISITION-CURVES FOR THE FOUR TRAINING GROUPS

(*S-100*, spaced-immediate; *S-50*, spaced-delayed; *M-100*, massed-immediate; *M-50*, massed-delayed)

running-times for Days 9 and 10 to logarithms and computing an analysis of variance. The results of this analysis, given in Table I, revealed that neither of the main effects, trial-distribution or condition of reinforcement, nor the interaction between them, was statistically significant.

The performance of the four training groups during extinction is shown in

TABLE I

ANALYSIS OF VARIANCE FOR THE LAST TWENTY TRAINING TRIALS

Source of variation	df.	Mean square	F
Rows (massed vs. spaced training)	1	.1071	1.31
Columns (immediate reward vs. delay)	1	.0114	—
R×C interaction	1	.2312	2.83
Within condition	44	.0817	
Total	47		

FIG. 2. The running times for each animal were converted to logarithms and an analysis of variance was performed. This analysis is summarized in Table II. It can be seen that only the reinforcement-condition had any effect upon extinction: *S*s which had experienced partially delayed reinforcement ran significantly faster

($P < 0.05$) on the extinction trials than did those who had been reinforced immediately. None of the other main effects nor their combinations led to significant variations in extinction-performance.

It will be noted in Fig. 2 that the groups were not at the same level of per-
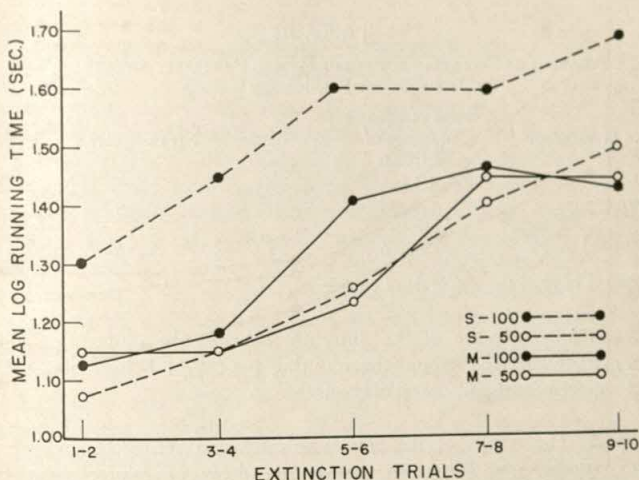
FIG. 2. EXTINCTION-CURVES FOR THE FOUR TRAINING GROUPS

(*S-100*, spaced-immediate; *S-50*, spaced-delayed; *M-100*, massed-immediate; *M-50*, massed-delayed)

formance at the termination of training, and it may be asked whether the difference in extinction between the partially delayed and immediately reinforced was related

TABLE II

ANALYSIS OF VARIANCE OF THE EXTINCTION-PERFORMANCE

Source of variation	df.	Mean square	F
Rows (massed vs. spaced training)	1	.1202	2.34
Columns (immediate reward vs. delay)	1	.2578	5.02*
Blocks (massed vs. spaced extinction)	1	.0768	1.49
R×C interaction	1	.1652	3.21
R×B interaction	1	.0334	—
B×C interaction	1	.0107	—
R×B×C interaction	1	.0037	—
Within Group	40	.0514	
Total	47		

* Significant beyond the 5% level of confidence.

to the different levels attained in acquisition. To answer this question, an analysis of covariance was carried out using the log-running-times on the last two days of acquisition as the concomitant variable and the log-extinction-scores as the criterion-variable. Since the assumptions of homogenous correlation coefficient and linearity of

regression could not be met for all the treatments in the factorial design, Groups
M-100 and *S-100* were pooled and tested against a pooled group consisting of
M-50 and *S-50*.[6] The summary of the analysis of covariance is presented in Table
III. Statistically controlling the level of performance at the end of training served

TABLE III

ANALYSIS OF COVARIANCE FOR THE POOLED PARTIALLY DELAYED AND
IMMEDIATELY REINFORCED GROUPS

Source of variation	Sums of squares of errors of estimate	df.	Mean square	F
Total	2.3294	46		
Within groups	2.0078	45	.0446	
Adjusted means	.3216	1	.3216	7.21*

* Significant beyond the 2% level of confidence.

to increase the significance of the difference between the immediately reinforced
and the partially delayed group, the resulting *F* of 7.21 being only slightly less
than that required for the 1% level of confidence.

Conclusion. The results of the present investigation verify the previous finding
of Crum, Brown, and Bitterman that partial delay of reinforcement produces
greater resistance to extinction than does immediate reinforcement. This effect,
which is unrelated to the distribution of training or extinction trials, constitutes a
problem for behavior-theory which may be just as formidable as the problem of
partial reinforcement.

---

[6] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education,* 1953, 323.

STIMULUS–VARIABILITY AND OPERANT DISCRIMINATION IN HUMAN SUBJECTS

By Edward J. Green, Dartmouth College

Statistical learning theory states that rate of conditioning is dependent upon the sampling probability, θ, of stimulus-elements within a given set.[1] In stimulus-discrimination, elements which are not common to Subsets SD_1 and SD_2 that are to be discriminated, and which have high θ-values, would facilitate discrimination, while elements which have low θ-values would lead to slower changes in behavior.

Two ways by which θ-values of such elements might be increased suggest themselves. If overlap between Subsets SD_1 and SD_2 is reduced, then the proportion of what might be called 'discriminable' elements within each subset is increased, thereby increasing the likelihood that such elements will be sampled by the organism if the subset is sampled at all. A second way is to retain considerable overlap between subsets, but to reduce the number of different arrays of elements within the subset. This is accomplished quite simply by repeating particular stimulus-patterns over a series of presentations. The extreme of this procedure would be to present only one pattern representing SD_1 and one pattern representing SD_2 over and over again. The θ-values of elements in this case would be maximal, θ-values decreasing as more of the possible arrays were included.

In a later development, the sequence of presentation of stimulus-patterns has been shown to control responding to a considerable degree.[2] It would seem that holding the sequence of stimuli constant for Ss in one group, while varying it from S to S in another group, would result in significantly less inter-S variability in the former group.

Although the two studies reported here were not originally conceived within the theoretical framework of the statistical learning model, it was thought useful to examine the results of these experiments in evaluating the considerations outlined above.

EXPERIMENT I

The first experiment was performed to determine whether Ss would discriminate more completely with decreased overlap between subsets of discriminative stimuli; or, in other words, whether Ss would discriminate more when the subsets were less similar.

Method. Instructions to the Ss and the apparatus have been described in detail elsewhere.[3] The method may be summarized briefly as follows: S sits in a chair

[1] W. K. Estes and C. J. Burke, A theory of stimulus-variability in learning, *Psychol. Rev.*, 60, 1953, 276-286.
[2] W. K. Estes and J. H. Straughan, Analysis of a verbal conditioning situation in terms of statistical learning theory, *J. exp. Psychol.*, 47, 1954, 225-234.
[3] E. J. Green, Concept formation: A problem in human operant conditioning, *J. exp. Psychol.*, 49, 1955, 175-180.

before a panel that separates him from *E*. A counter is mounted beside an opening in the panel, and a telegraph key is mounted on a table before the panel.

S is told that a series of cards will be shown to him, and that he is to tap the key whenever he thinks a 'right' card (SD_1) is being shown. He is also told that when a 'right 'card is being shown he will score points on the counter for tapping the key and that he may get more than one point during the time a 'right' card is being shown, so he must tap the key repeatedly if he is to get all the points possible. The *S* is able to note the operation of the counter. He is also told that he will get no points for tapping the key while a 'wrong' card (SD_2) is being shown. He is to try to find out what it is about a card that makes it 'right,' and also to try to get as high a score as he can. A summary of the instructions is attached to the panel so he may refer to them during the course of the experiment.

Stimulus-cards are then presented serially, 1 min. each, through the opening in the panel, and *S* scores points for tapping the key on appropriate occasions. A series of 20 cards is used, half SD_1 and half SD_2, during conditioning of the discrimination; these are followed by a second series of 20 cards, half SD_1 and half SD_2, presented without interruption during extinction.

Reinforcement was accomplished by crediting *S*s with points on the counter on a 15:1 fixed-ratio schedule.

Five *S*s were run in each of the three groups. The experimental group was presented with a set of discriminative stimuli in which no elements were common to both SD_1 and SD_2. Stimulus-patterns were made up of a matrix of nine circles, represented by letters in the following manner:

$$a \quad b \quad c$$
$$d \quad e \quad f$$
$$g \quad h \quad i$$

In a reference group, SD_1 was characterized by having circles *a, b,* and *c,* blacked-in. SD_2-patterns might have none, any one, or two of these three circles blacked-in, but never all three. For every instance of SD_1, with a particular subgroup of the remaining six circles filled in, there was a corresponding SD_2 in the series with the same subgroup filled in. There were no duplicated patterns in the 40-card series.

For the experimental group, a series of cards was prepared where, in addition to circles *a, b,* and *c,* only three circles in the lower six positions, Circles *d, e,* and *g,* were blacked-in for SD_1. The only circles blacked-in for SD_2 were Circles *f, h,* and *i.* It was necessary to duplicate patterns in this series, since only eight different patterns are possible. A 40-card series was used as before.

The third group was established as a control to determine that effects obtained with the new stimulus-series were not due merely to repetition of patterns rather than to dissimilarity between SD_1 and SD_2. Eight SD_1 and eight corresponding SD_2-patterns were selected at random from the stimulus-series of the reference group. These patterns were duplicated as necessary to fill out a third 40-card series which was presented to *S*s of the control group.

Results. The mean numbers of key-tapping responses are presented in Table I. The data were normalized by the log transform, and the difference in responding to SD_1 and SD_2 was tested by the two-tailed *t*-test in the three groups. Chi square

for the summed probabilities of the ts so obtained in conditioning and extinction was significant beyond the 0.001 level of confidence, indicating that satisfactory discriminations were developed. The difference between number of SD_1 and number of SD_2 responses was taken as an index of the degree of discrimination. Chi square was used to compare the three groups in this respect, and was found to be significant beyond the 0.001 level in both conditioning and extinction.

The greatest difference in responding to SD_1 and SD_2 during conditioning was

TABLE I

MEAN NUMBER OF KEY-TAPPING RESPONSES IN CONDITIONING AND EXTINCTION
(Experiment I)

Group	Conditioning		Extinction	
	SD_1	SD_2	SD_1	SD_2
Reference	2585.8	708.8	1171.2	822.8
Experimental	3174.2	201.2	325.0	215.6
Control	3143.4	414.0	440.0	336.0

found in Ss of the experimental group. The next greatest difference was found in Ss of the control group. In other words, repetition was found to favor discrimination, but the use of dissimilar stimulus-classes was found to facilitate it even more.

Resistance to extinction was greatest in the reference group, less in the experimental group, and least in the control group.

EXPERIMENT II

In the previous experiment, the stimulus-cards were thoroughly shuffled for each S. It was thought that this procedure might have introduced additional inter-S variability. The present experiment was conducted to determine whether inter-S variability might not be reduced by presenting each S with the same sequence of stimulus-cards.

Method. The reference group of Experiment I also served as a reference group in this experiment. An additional group of five Ss was used. The procedure for this latter group differed from that of the reference group only in that the same sequence of discriminative stimuli was employed with all Ss.

Results. The number of responses under the various conditions was normalized by the log transform. F-ratios between the reference and experimental groups are presented in Table II. The total numbers of responses in SD_1 were not significantly more or less variable in either group; but SD_2 responding was significantly less variable in the experimental group than it was in the reference group.

DISCUSSION

Ss in Experiment I discriminated between the dissimilar stimulus-classes more readily than did Ss who were required to discriminate between stimulus-classes having common elements as specified in this paper. Repetition of stimulus-patterns hastened discrimination, but not as much as did repetition plus dissimilarity, or decreased overlap of stimulus-elements between SD_1 and SD_2. This would seem

to support the considerations growing out of the learning model. θ-values were increased by repetition in the control group of Experiment I, but they were increased even more by the decrease in overlap between SD_1 and SD_2 in the experimental group.

It seems that rate of responding is primarily under the control of the schedule of reinforcement once responding has begun in the presence of a particular stimulus-pattern, since responding to SD_1 did not seem particularly sensitive to the condi-

TABLE II

F-Ratios of Variances of log R (Key-Tapping Responses)
(Experiment II)

Series	Stimulus	F	p
Conditioning	SD_1	1.82	>0.05
	SD_2	8.18	<0.05
Extinction	SD_1	1.98	>0.05
	SD_2	10.99	<0.05

tions of the experiment. Most of the effect of both repeating stimulus-patterns and using dissimilar subsets is seen in responding to SD_2.

It seems clear that specifying and manipulating stimulus-elements as was done in this experiment has operational utility. Certainly the proportion of the total population of elements subject to experimental manipulation may be small, yet the response selected is sensitive to changes such as were made in these experiments.

According to the model, resistance to extinction should have been greatest in the reference group of Experiment I; less for the control group, and least for the experimental group.

If total numbers of responses emitted during extinction are used as a measure of resistance to extinction, this is what obtained. If, however, the difference between numbers of responses in SD_1 and SD_2 is used, as was originally intended, as the measure of degree of discrimination, a reversal is found between the predicted and obtained results in the experimental and control groups. No explanation is advanced here to account for this reversal.

Experiment II demonstrated that holding the sequence of stimuli constant from S to S did decrease inter-S variability in responding, which indicates that the sequence of stimuli is an important parameter. The greatest reduction in variability brought about by the introduction of the single sequence was observed in responding to SD_2. Again the particular insensitivity of responding to SD_1 and the sensitivity of responding to SD_2 raise interesting questions about the effects of using schedules of reinforcement in developing discrimination.

The original procedure of shuffling stimulus-cards before they were shown to S was developed to avoid artifacts in discrimination arising from any given sequence of presentation. Certainly the sequence used in this experiment, or any particular sequence, might be expected to produce artifacts; but the reduction of inter-S variability seems to balance the undesirability of such artifacts, especially, since they can easily be isolated and controlled.

CONCLUSIONS

Increasing θ-values, either by reducing overlap between subsets of discriminative stimuli, or by repeating particular stimulus-patterns, has been shown to facilitate discrimination. Ss more readily discriminated between subsets of stimuli when there was less overlap between subsets; that is, when the subsets were less similar. Repetition of stimulus-patterns facilitated discrimination, but the facilitation was not so great as when, in addition, the overlap between SD_1 and SD_2 was also less.

Inter-S variability of key-tapping responses was reduced significantly in SD_2 during both conditioning and extinction when the same sequence of discriminative stimuli was presented to each S. Variability in responding to SD_1 was not significantly altered by this means.

COMPARISONS OF INCIDENTAL AND INTENTIONAL LEARNING WITH DIFFERENT ORIENTING TASKS

By Irving J. Saltzman, Indiana University

Previous comparisons of incidental and intentional learning have shown that the superiority of intentional learning depends in part on whether the intentional learners are required to perform an orienting task.[1] Neimark and Saltzman have described conditions under which intentional learners made higher scores only when they did not perform such a task.[2] The present study was designed to determine whether the difference between incidental and intentional learning might be influenced by variation in the nature of the orienting task.

In the experiment of Neimark and Saltzman, the learning material was a list of 14 two-digit numbers and the orienting task was to find and circle on a prepared sheet each of the numbers as it appeared in the window of a memory-drum. When the rate of presentation was decreased from 3 sec. to 6 sec. per number, the scores of intentional learners were reliably superior even though they were required to perform the task. In the present study, incidental and intentional learning were compared under conditions identical with those of the earlier experiment, but with different groups performing four additional tasks. The new tasks were designed to eliminate the superiority of intentional learning, an end which might be accomplished, it was believed, in either of two ways. Three of the new tasks were expected on rational grounds to *facilitate incidental learning,* while the other was expected to *interfere with intentional learning.*

Procedure. The Ss were 200 undergraduate students at Indiana University. They were divided into 10 groups of 20 Ss each; 5 groups learned incidentally and 5 intentionally. The learning material for all groups was the same: a list of 14 two-digit numbers. The Ss, studied individually, were seated at a table on which was located an electrically operated memory-drum. The drum presented the numbers at the rate of one number every 6 sec. The list of numbers was presented four times, each time in a different order, with a 10-sec. interval between successive presentations. The Ss in all 10 groups were required to perform orienting tasks, but only the Ss in the five intentional groups, of course, were informed that there would be a test for the numbers and told to try to learn them. Each of the five different orienting tasks was performed by one intentional and one incidental group. The tasks were as follows:

Task 1. The Ss were told that as each number appeared in the aperture of the memory-drum, they were to look for and circle the number on a prepared number-

sheet which contained all of the numberrs from 11 to 99 arranged in sequence in a systematic order. A different sheet, with the numbers arranged in a different order, was used on each of the four trials. This was the orienting task used by Neimark and Saltzman, and with which intentional learning was found to be superior to incidental learning (when the 6-sec. rate was employed).

Task 2. The Ss were told that as each number appeared in the aperture of the memory-drum they were to look for and circle the number on the prepared sheet, then, to add 5 to the number mentally, and finally, to write the total, the number *plus 5,* on a pad of paper, twice. A different sheet with the numbers arranged differently was used on each of the four trials.

Task 3. The Ss were told that as each number appeared in the aperture of the memory-drum, they were to call the number aloud, find it and circle it on the number-sheet, and, finally, write the number three times on a pad of paper, saying it aloud each time. A different sheet with the numbers arranged differently was used on each of the four trials.

Task 4. The Ss were told that on the first trial, as soon as they noticed a number in the aperture of the memory-drum, they were to close their eyes and say the number over and over to themselves until they heard the click of the apparatus indicating that the next number had appeared. On the second trial, they were told to do the same thing, but in addition to try to visualize the number while saying it. On the third trial, as each number appeared in the aperture of the memory-drum, the Ss were told to imagine writing the number again and again on an imaginary piece of paper until the next number appeared. On the fourth trial, the Ss were told to do all three together, say the number over and over silently, visualize it, and imagine writing it repeatedly on an imaginary piece of paper.

Task 5. As each number appeared in the aperture of the memory-drum, the Ss were told that they were merely to notice whether or not the different numbers suggested anything to them, such as birth-dates, street-addresses, telephone-numbers, and the like. They were told that they would be asked about their associations after four presentations of the numbers.

None of the Ss in the *incidental* groups was told, either specifically, or by implication, to learn the numbers. The incidental learners who performed Tasks 1, 2, and 3 were told that E was interested in studying the different ways of arranging the numbers on the number-sheets in preparation for some later work on the decoding of number-messages. The incidental learners who performed Task 4 were told that E was interested in ESP and that he wished to compare the effectiveness of four different methods of sending telepathic messages. The incidental learners who performed Task 5 were told that E was interested in the range of associations evoked by numbers. All of the Ss in the intentional learning groups were told that E wished to find out how fast they could learn the numbers under some special learning conditions. The special learning conditions were the performances of the various orienting tasks while trying to learn the numbers.

After the fourth trial, all Ss were given recognition-tests for the numbers. The 14 numbers of the learning series plus 42 additional two-digit numbers were randomly arranged in four columns on a mimeographed sheet. Each S was requested to circle all the numbers which he recognized as numbers that had appeared on the memory-drum, 2 min. being allowed for the test. Since all of the Ss did not

circle the same number of items, a correcting procedure was used in scoring the recognition-test. Scores on the test were derived by subtracting from the number of correctly circled items the number that might be expected to be correct by chance. For every four numbers circled, one number was considered to be correct by chance. For example, if S circled 12 numbers, 10 of which were numbers that had been on the list, his score would be seven: 10 *minus* one-fourth of 12.

Results. An analysis of variance yielded significant F-values for the differences between the orienting tasks ($P = 0.01$), for the difference between incidental and intentional learning ($P = 0.01$), and for the interaction ($P = 0.05$). Means of the derived scores, and the *t*-ratios and probabilities obtained from comparisons of the

TABLE I

MEANS AND DIFFERENCES BETWEEN MEANS OF DERIVED SCORES FOR INTENTIONAL
AND INCIDENTAL LEARNING WITH DIFFERENT ORIENTING TASKS

Orienting task	Condition		Diff.	t	P
	intentional	incidental			
1	6.41	4.26	2.15	3.47	.01
2	5.14	4.97	0.17	0.27	.80
3	6.32	5.64	0.68	1.10	.30
4	6.89	4.94	1.95	3.14	.01
5	7.60	7.46	0.14	0.22	.90

pairs of incidental and intentional groups, are shown in Table I. It can be seen 'hat with two of the orienting tasks, Tasks 1 and 4, intentional learning was reliably superior to incidental learning, but that with the remaining three tasks, Tasks 2, 3, and 5, the differences between the means were not reliable. These results, in light of the significant interaction, indicate that the superiority of intentional over incidental learning is related to the nature of the orienting task. A reliable difference may be found between incidental and intentional learning with one orienting task but not with another.

It will be recalled that in selecting the orienting tasks for this study an attempt was made to choose tasks which might eliminate the difference between incidental and intentional learning found in the earlier study by Neimark and Saltzman. Two different kinds of orienting task were used, each kind being expected to accomplish the same end in a different way. One kind of orienting task was expected to interfere with the effort to learn in the intentional group, while the second was expected to facilitate incidental learning. Task 2, which required the Ss to add 5 to each number and to write the total twice, was expected to interfere with intentional learning, while Tasks 3, 4, and 5 were expected to facilitate incidental learning. Task 1, of course, was known from the study of Neimark and Saltzman to make for better learning under intentional as compared with incidental conditions. Tasks 2, 3, 4 and 5 were selected on purely intuitive grounds. It was believed that when the intentional learners performed Task 2, they would make lower scores than when they performed Task 1 because (a) there would be less time left after performing Task 2 in which to try to learn than there was after performing Task 1, and (b) adding 5 to the numbers and writing the total twice would introduce an opportunity, not afforded by Task 1, for confusing the original numbers with the new numbers. It was believed that with Task 3, 4, or 5 the incidental learners would

make higher scores than they did with Task 1 because each of these tasks was designed to make the incidental learners do something that intentional learners presumably do ordinarily when they are trying to learn, *e.g.* say the numbers repeatedly, write them down, associate familiar things with them.

With three of the four tasks the expected results were obtained. The scores of the intentional learners who performed Task 2 were reliably lower than those of the intentional learners who performed Task 1 ($P = 0.05$). The corresponding difference between the scores of the incidental learners on Tasks 1 and 2 was not significant ($P = 0.30$). Because of this differential effect of the orienting tasks on incidental and intentional learning, the superiority of intentional learning over incidental learning found for Task 1 was not found for Task 2.

The performance of Task 3, 4, or 5 was expected to promote more learning by the incidental learners than did their performance of Task 1. This expectation was supported by the results obtained with Tasks 3 and 5, but not with Task 4. The incidental learners who performed either Tasks 3 or 5 made reliably higher scores than the incidental learners who performed Task 1 ($P = 0.05$ and 0.01, respectively); the incidental learners who performed Task 4 did not make reliably higher scores than the incidental learners who performed Task 1 ($P = 0.30$). Changing from Task 1 to any of these tasks had no effect on the scores of the intentional learners ($P = 0.90$, 0.50, and 0.10 for Tasks 3, 4, and 5, respectively).

In previous studies, the interfering effect of orienting tasks on intentional learning was suggested and demonstrated.[3] The results of the present study, with Task 2, confirm the earlier findings. In addition, the present study reveals another kind of influence of the orienting task—a facilitating effect on incidental learning. Although the consequence of these two effects is the same—the elimination of the superiority of intentional over incidental learning—the existence of the two processes should be recognized.

---

[3] Saltzman, The orienting task in incidental and intentional learning, this JOURNAL, 66, 1953, 593-597; Neimark and Saltzman, *op. cit.*, 618-621.

THRESHOLD-LUMINANCE FOR RECOGNITION IN RELATION TO
FREQUENCY OF PRIOR EXPOSURE

By KATHERINE E. BAKER and HERMAN FELDMAN, University of Nebraska

The present study is concerned with the effect of familiarity on the visual recognition of nonsense-words. Solomon and Postman have shown that duration of exposure required for correct recognition of such words is a negatively accelerated function of the number of prior exposures of the words in a preliminary session.[1] Their results for a second measure—'trials'—roughly indicate that the relationship is similar when threshold is measured in terms of luminance, but their method, like that employed in other related experiments,[2] was crude and did not permit expression of thresholds in terms of photometric brightness. The relation of luminance-threshold to familiarity with nonsense-words is here investigated under more appropriate conditions.

Procedure. Thresholds were measured in 12 undergraduate students immediately following 1, 2, 5, 10, or 25 exposures to each nonsense-word of 'a list of 10 employed by Solomon and Postman. Essentially the same but some shortened instructions were given, leading each *O* to believe that the experim concerned foreign-language training. Each *O* pronounced, one at a time, the 10 core-words and 14 other filter-words as they appeared randomly in a deck of 100 shuffled cards containing duplicates of the core-words in appropriate numbers. For different *Os* the frequencies of the latter words were systematically varied to counterbalance intrinsic differences in recognizability.

For the measurement of recognition-thresholds, *O* was seated 8 ft. from a 35 × 30 in. sandblasted glass screen in a completely dark room, and the test-words were projected onto the back of the screen from an adjoining room. The *O* was instructed to keep his eyes directed at the screen, especially following each ready-signal, and to report after each flash on the screen. A response of 'No' was to be given if the word was not recognized; otherwise *O* was to say what he thought the word was, using an intercommunication system between the two rooms. The procedure of Solomon and Postman was followed with respect to practice-words, order of presentation, and the inclusion of 10 English and 10 extra nonsense-words among the core-words during threshold-tests.

The projection-system consisted of a modified Viewlex slide-projector situated 6 ft. from the glass screen and directly in line with the *O* sitting in the adjoining room. Stimulus-words, on 2 × 2 in. transparent slides, were so focused on the screen that the typewritten letters measured $1\frac{1}{8} \times 1\frac{1}{2}$ in. To obtain the proper

[1] R. L. Solomon and Leo Postman, Frequency of usage as a determinant of recognition thresholds for words, *J. exp. Psychol.,* 43, 1952, 195-201.

[2] J. C. Gilchrist and L. S. Nesberg, Need and perceptual change in need-related objects, *J. exp. Psychol.,* 44, 1952, 369-376; J. C. Gilchrist, J. F. Ludeman, and W. Lysak, Values as determinants of word recognition thresholds, *J. abnorm. & soc. Psychol.,* 49, 1954, 423-426.

range of luminance, the voltage across the 150-w. lamp of the projector was reduced, and a Corning glass daylight-filter (No. 5900) was mounted immediately in front of the lens. A piece of ¼-in. milk glass placed just in front of the projector-lamp served to diffuse the light evenly over the screen. The projector was fitted with a No. 4 Ilex Universal timing-shutter which was set for an exposure of 0.2 sec. on all trials.

Control of luminance-level was accomplished by means of polaroid disks contained in a special mounting placed immediately in front of the lens and daylight-

TABLE I

THRESHOLD-LUMINANCE (FT.-L.) FOR RECOGNITION OF NONSENSE-WORDS WITH DIFFERENT FREQUENCIES OF PRIOR EXPOSURE

Frequency	Mean log-L	SD log-L	Geom. mean-L
1	−0.883	0.030	0.1309
2	−0.899	0.037	0.1262
5	−0.903	0.030	0.1250
10	−0.923	0.041	0.1194
25	−0.974	0.033	0.1062

filter. The angular displacement of the axes of the rotating disk and the stationary disk could be read to the nearest 1° from a calibrated scale. The maximal luminance was measured to be 0.1425 ft.-L with a Macbeth Illuminometer placed on O's side of the screen. Other luminance-levels less than the maximum were found from the relationship between the angle of displacement of polaroids and measured luminance values obtained at 5° intervals over the scale. Care was taken to avoid very low settings where color-changes are likely to occur.

For the threshold-determinations the polaroid disks were set to provide low luminance, a ready signal was given, and then the word was flashed onto the screen for 0.2 sec. Luminance-level was increased in approximately 0.005 log-unit steps on succeeding trials by changing angular displacement until correct recognition was achieved. The threshold for each word was measured in one such ascending series for each O.

Results. The raw data, in terms of angular displacement of the polaroids just great enough for correct recognition, were utilized to find the logarithm of threshold-luminance for each threshold-determination. The threshold-luminances in logarithmic form were averaged for the 12 Os for the two words representing each of the five frequencies of prior exposure. The resultant mean log-luminance levels required for correct recognition are presented in Table I along with the standard deviations. To allow for reference of the data to the foot-lambert scale, the geometric mean-luminances are included in the table.

As may be seen in Fig. 1, threshold-luminance is a decreasing function of the frequency of prior exposure. The change in threshold represents about a 19% reduction over the range of frequencies used and the curve is only slightly negatively accelerated.

An analysis of variance was performed on log-luminance measures for the core-words. Data in logarithmic form were used in order to increase normality and homogeneity of the distributions. The analysis yielded an F significant beyond the 1% level of confidence for the effects of frequency of prior exposure. A significant difference at better than the 1% level of confidence was indicated by a t-test

between frequencies of exposure of 10 and 25; thus it would appear that the curve is still dropping and the minimal threshold has not yet been reached at 10 exposures.

These data attest that luminance-level required for correct recognition is an acceptable type of measurement to make in perceptual experiments. The relationship between recognition-threshold and frequency of prior exposures is of the same general kind as that found by Solomon and Postman for duration. There are, however, certain differences in the two sets of findings which should be noted.
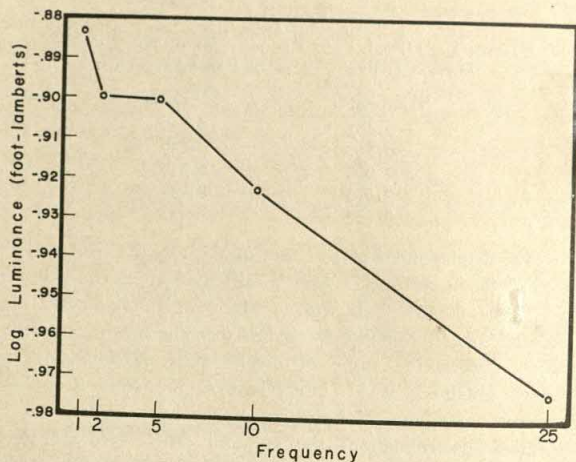
FIG. 1. LOG THRESHOLD-LUMINANCE (FT-L.) FOR CORRECT RECOGNITION OF NONSENSE-WORDS WITH DIFFERENT FREQUENCIES OF PRIOR EXPOSURE

Time-thresholds were reduced by a larger amount (70%) over the comparable range of frequencies of prior exposure. Since the smaller change in luminance-thresholds is statistically significant, it seems that there is a difference between luminance and time with respect to visual functioning. The function relating luminance-thresholds to frequency of prior exposure is less accelerated and appears still to be dropping at 25 exposures, at which point the analogous function for time seems to have reached an asymptote. Aside from these differences, which may be of some importance in certain comparisons of experimental findings, the two techniques may be considered comparable, at least for the recognition of nonsense-words. The present test of the equivalence of the two measures appears to be more convincing than that of Solomon and Postman who used number of trials to represent unspecified luminance-levels. Both experiments demonstrate that frequency of prior exposure is a factor in determining perceptual thresholds.

Summary. Each of 10 nonsense-words was exposed 1, 2, 5, 10, or 25 times to 12 *O*s, and recognition-thresholds were then measured in terms of the minimal luminance-level required for correct identification. Threshold-luminance, like threshold-duration, is a function of familiarity, but there seem to be differences in the two functions.

THE EFFECT OF INSTRUCTION AND DEGREE OF TRAINING ON SHIFTS OF DISCRIMINATIVE RESPONSES

By Harold W. Stevenson and George Moushegian, University of Texas

The effects of varying numbers of reinforcements on the ability of human adult Ss to shift from one discriminative response to another have been investigated in several recent experiments. Grant and his colleagues found that the ease of shifting a response in the Wisconsin Card-Sorting Tests was positively related to the degree to which the responses were reinforced.[1] Ss who were required to sort the cards 10, 20, or 40 successive times according to one method were able to switch to a new method more easily than were Ss who were required to make fewer sortings. In problems where Ss learned sequences of lever-pressing responses, Mandler found zero or slight negative transfer when 10 to 100 successive correct responses were required on one problem before a new problem was introduced.[2] Vogt, in a study of discriminative learning, found that learning to discriminate position was retarded after practice in discriminating size, but that the degree of retardation did not differ when the criterion of the original learning had been 4, 8, or 12 successive correct size-responses.[3]

In accounting for the results, several hypotheses have been advanced. Mandler used the principles of reinforcement in a theoretical discussion of transfer.[4] An interpretation in terms of learning set was offered by Grant, who assumes that greater reinforcement makes it easier for an S to learn how to learn. Another suggestion made by Grant is that extensive reinforcement makes the correct response "stand out so that S may more readily react to a shift or a change."[5] The poorer performance of the Ss given minimal training would be due in part, therefore, to their difficulty in discerning when one problem ended and another began.

The purpose of the present experiment was to determine the effects of varying the amount of over-training over a broad range, and to discover whether the poor performance previously found in Ss given minimal training could be improved by giving them explicit instructions concerning the change in problems.

Procedure. The Ss were 108 volunteers from the introductory classes in psychology. The apparatus consisted of: (a) three 3-in. square blocks, each bearing a white

---

[1] D. A. Grant and E. A. Berg, A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem, *J. exp. Psychol.,* 38, 1948, 404-411; D. A. Grant and J. R. Cost, Continuities and discontinuities in conceptual behavior in a card sorting problem, *J. gen. Psychol.,* 50, 1954, 237-244.

[2] George Mandler, Transfer of training as a function of degree of response overlearning, *J. exp. Psychol.,* 47, 1954, 411-417.

[3] R. D. Vogt, The effect of overlearning on change of set, unpublished Masters thesis, University of Texas, 1953.

[4] Mandler, Response factors in human learning, *Psychol. Rev.,* 61, 1954, 235-244.

[5] Grant and Cost, *op. cit.,* 242.

cardboard square of different size; (b) a tray for presenting the blocks; (c) a screen to shield E; and (d) a rubber eraser which served as the goal-object. The areas of the squares were ¾, 1, and 1-½ sq. in. On the tray were three wells in which the goal-object could be concealed. The tray and blocks were painted flat black.

The Ss were tested individually. S was seated across from E who demonstrated how the blocks could be used to cover the wells, and then instructed him as follows:

These blocks fit over these holes in this way. Now, see this (eraser) ? I am going to hide it and I want you to find it as often and as consistently as you can. The problem is solvable.

Following this, E arranged the blocks over the proper wells for the first test-trial. The screen was removed and S attempted to find the test-object. After S made a choice the screen was again interposed and the blocks rearranged. The arrangement of the blocks on successive trials followed a randomized schedule, but was the same for all Ss.

Every S was required to solve two problems. The correct response for Problem 1 was a size-response, and the correct response for Problem II was a position-response, *i.e.* the eraser was placed under a particular block in Problem I and in a particular well in Problem II. There were two further experimental variables: (a) the criterion to which Problem I was learned; and (b) instructions to one group concerning the change in problems.

Three groups of Ss were trained to criteria of 4, 16, or 40 successive correct responses on Problem I. On the trial following this training, Problem II was introduced immediately and without comment. For these groups there were no additional cues that the problem had changed. Instructions concerning the change were given to a fourth group of Ss trained to the four-trial criterion. These Ss were instructed after solving the first problem; "Now the problem will change, and you are to find a new solution." The four groups may be designated Groups 4, 16, 40, and 4ᵢ, according to learning criterion and the presence of instructions.

The criterion for solution of Problem II was four successive correct responses. Three Ss were tested on each of the 9 block-well combinations, making a total of 27 Ss in each group. At the conclusion of the experiment the Ss were asked to give a detailed verbal report concerning the bases of their responses.

Results. The mean numbers of trials required for Problem I and II are presented in Table I. The numberrs are exclusive of the criterion-trials. The distributions of trials for both problems were highly skewed, hence the nonparametric F-test was used in the analysis of the data.[6]

Comparison of the means for Problem I reveals that none of the differences is statistically significant. This serves as a control on the equality of the groups. The speed with which Problem II was learned did differ, depending upon the amounts of practice given on Problem I. Group 4 required nearly twice as many trials to solve the problem as did Group 16 ($F = 1.96$, $P < 0.001$). Group 16, in turn, required somewhat fewer trials than did Group 40, but the difference is not significant ($F = 1.31$, $P > .05$). We may assume that there is little difference

---

[6] Leon Festinger, A statistical test for means of samples derived from skew populations, *Psychometrika*, 8, 1943, 205-210.

in speed of learning following a 16- to a 40-trial criterion, but that the use of a lower criterion produces slower learning.

The instructions given Group 4I increased the speed of learning on Problem II over that of Group 4, which received identical training except for the instructions.

TABLE I

THE MEAN NUMBERS OF TRIALS REQUIRED TO LEARN PROBLEMS I AND II

Group	Problem I		Problem II	
	M	SD	M	SD
4	11.4	15.9	22.5	13.6
16	10.3	12.7	11.5	7.9
40	11.4	14.3	15.1	14.7
4I	8.3	10.5	11.0	10.5

The difference is significant beyond the 0.001 level of confidence $(F = 2.04)$.

A breakdown of the errors made in solving Problem II is presented in Table II. An error is considered to be perseverative when the block chosen is the one that was correct on Problem I. The proportion of errors that were perseverative is quite constant for Groups 4, 16, and 40, but there is a notably smaller proportion

TABLE II

MEAN NUMBERS OF PERSEVERATIVE AND NON-PERSEVERATIVE ERRORS FOR PROBLEM II

Group	Perseverative	Non-perseverative	% Perseverative
4	6.1	8.6	42
16	3.2	4.6	39
40	4.4	6.1	42
4I	1.8	5.1	30

of perseverative errorrs in Group 4I. The statistical significance of this difference may be determined by computing the proportion of perseverative errors for each S in Group 4 and 4I and ascertaining whether the means of these proportions differ significantly. The difference obtained by this procedure is significant between the 1% and 5% levels of confidence $(F = 1.68)$.

Discussion. The results provide further evidence that in situations involving the same stimulus-objects, small amounts of training may result in greater difficulty in solving successive problems than larger amounts of training. The Ss' behavior and verbal reports revealed several reasons for this. The improved learning following instructions indicates that one of the difficulties of switching responses after minimal practice may be, as Grant has suggested, that S finds it difficult to discern that a switch has occurred and that he must change responses. The reports of the Ss in Group 4 were very confused and some Ss were unaware that they had solved a problem. In contrast, all of the Ss in Groups 16 and 40 were able to give accurate descriptions of the solutions to both problems. Extended practice may also provide useful clues about the level of difficulty of later problems. This would be especially helpful in the present experiment, for the initial hypothesis stated by many Ss was that the correct response depended on patterns of blocks or sequences of positions. Such complex hypotheses obviously had to be rejected before the simple, correct

response could be found. Finally, the Ss in Groups 16 and 40 reported that they anticipated the switch and tried to determine the types of responses that could be required. Because of their greater time for practice, these Ss were able to begin solving the second problem while they were still responding to the first.

Summary. Three groups of 27 Ss were trained to make a size-discrimination. The problem was then changed, without their knowledge, to a position-discrimination. A fourth group was instructed that the problem would change. The results indicate (a) that the degree of training in the first discrimination was positively related to ease of solving the second; (b) that perseveration of responses was not a function of the degree of training; and (c) that instructions significantly increased the ease of making a change.

ESTHETIC FATIGUE IN RANKING

By EVELYN M. WEST, Columbia University, and
A. W. BENDIG, University of Pittsburgh

Both Guilford and Woodsworth have listed as one of the major disadvantages of the paired comparisons and ranking methods, the esthetic fatigue generated by requiring the S to make a large number of judgments.[1] Although Hamlin and Cummings have used this construct to explain some of their results on clinical judgment, the existence of this judgmental phenomenon has not been empirically demonstrated.[2] From these discussions it is somewhat difficult to conclude how judgments are affected by 'esthetic fatigue,' but apparently later judgments are more subject to error-variance, and the amount of error in the judgments is a monotonically increasing function of the number of preceding judgments elicited from the S. This increased error-variance, if random, should decrease both the reliability and validity of judgments that are subject to more 'esthetic fatigue.' Several methods might be used to test this hypothesis: (a) the group, or inter-judge, reliability of the first and last judgments elicited from Ss may be compared, (b) the expected decline in the validity of later judgments of a single S may be assessed where an outside criterion is available, or (c) the test-retest, or intra-judge, reliability of individual judges can be compared, where varying numbers of similar judgments are interpolated between the first and second judgments of the same stimuli. Bendig suggested that 'esthetic fatigue' is one explanation of the differences found in the test-retest reliabilities of the first and second sets of esthetic stimuli judged by his Ss.[3]

The study reported here again used method (c) above. It was designed to test the hypothesis of the existence of 'esthetic fatigue' in a judmental situation.

Procedure. A set of 96 small (1¼ × 2 in.) colored paintings was mounted individually on white cards (3 × 5 in.) for convenience in handling. Five packs, each containing 12 cards, were assembled by randomly drawing from this larger set.[4] The Ss were given one or more of the shuffled packs and asked, after they had an opportunity to inspect the paintings in the pack, to rank the 12 stimuli as to esthetic merit, from the one they liked best to the one they liked least. All Ss were tested individually and their ranking was done under the illumination of a daylight 15-w. fluorescent lamp.

The total sample of 40 Ss was divided into a sample of 15 graduate students in

[1] J. P. Guilford, *Psychometric Methods*, 1936, 280; R. S. Woodworth, *Experimental Psychology*, 1938, 378.
[2] R. M. Hamlin, The clinician as judge: Implications of a series of studies, *J. consult. Psychol.*, 18, 1954, 233-238; S. T. Cummings, The clinician as judge: Judgments of adjustment from Rorschach single-card performance, *ibid.*, 243-247.
[3] A. W. Bendig, Inter-judge vs. intra-judge reliability in the order-of-merit method, this JOURNAL, 65, 1952, 84-88.
[4] Bendig, *op. cit.*, 85.

psychology (Group A) and a sample of 25 undergraduate students enrolled in a course in general psychology (Group B). The Ss of Group A were used to test the packs for differences in retest reliability as a preliminary step, while those of Group B comprised the 'esthetic fatigue' subgroup. Each S in Group A ranked one of the 5 packs, with three Ss randomly assigned to each pack, and re-ranked the same pack with a minimum of 20 hr. intervening between the sessions. The two

TABLE I

ANALYSIS OF VARIANCE OF TRANSFORMED (r-TO-z) RETEST RELIABILITY COEFFICIENTS FROM FIVE ESTHETIC FATIGUE SUBGROUPS

Source-variation	df.	Mean square	F
Total	24		
Packs	4	.196	1.32
Subgroups	4	.731	4.94*
Linear Regression	1	2.861	19.33†
Deviations	3	.021	.14
Error (P×S)	16	.148	

* Significant at the .01 level.
† Significant at the .001 level.

sets of ranks for each S were correlated by the usual rank-difference formula as a measure of intra-judge or retest reliability.

The 25 Ss of Group B were randomly divided into 5 subgroups of 5 Ss each. Each S ranked one of the 5 packs at the beginning of the experimental session (which lasted from 19 to 23 min.) and again at the end of the session with the 5 Ss in each subgroup ranking a different pack. The subgroups are differentiated by the amount and kind of intervening activity between their first and second rankings of the same pack. Subgroup 1 spent the intervening period in the non-esthetic activity of alternate trials at squeezing a hand dynamometer and working on a series of algebra problems. Subgroup 2 also had alternate dynamometer and algebra trials, but also ranked one of the other packs of esthetic stimuli during this period. Subgroup 3 ranked two of the other packs during this period (in addition to a lessened number of dynamometer and algebra trials), while Subgroup 4 ranked three other packs. Subgroup 5 spent virtually the entire time intervening between the first and second rankings of the same pack in ranking esthetically the remaining four packs of cards, with a different order of the four packs used with each S. As for Group A, the measure of intra-judge reliability for each S was the rank-difference correlation between his first and second rankings of the same pack.

Results. The rank-difference coefficient of correlation for each of the Ss in Groups A and B was transformed into a normally distributed variate by the usual r-to-z method.[5] The first step in the analysis was to test the hypothesis of similar retest reliabilities for the five packs using the data of Group A. The 15 z-measures from Group A were subjected to a simple analysis of variance and the variance between packs was found not to be significantly larger than the error-term. ($F = 1.08$

---

[5] Hartley has shown this procedure to be applicable to rank-difference coefficients of correlation. See H. O. Hartley, Approximate tests for comparisons of rank correlations. *J. Amer. statis. Assoc.*, 49, 1954, 329.

with 4 and 10 degrees of freedom). Apparently the randomization procedure used in assembling the packs had effectively equalized them in retest reliability.

The 25 zs from Group B were analyzed as a two-variable design. The results are given in Table I. The variance for packs again was insignificant, confirming the results from Group A on the homogeneity of the packs. The differences between subgroups, however, was significant at the 1% level, and the hypothesis of a linear

TABLE II

RETEST RELIABILITY IN ESTHETIC RANKING AS A FUNCTION OF AMOUNT OF INTER-VENING ESTHETIC JUDGMENTS

Subgroups	Average retest reliability	
	Mean z	Mean r
1	1.848	.952
2	1.734	.940
3	1.350	.874
4	1.178	.827
5	.930	.731

relation between retest reliability and amount of esthetic activity intervening between first and second rankings was confirmed at the 1/10% level of confidence. Table II gives the average z and equivalent r for each of the five subgroups, and it can be noted that the average z decreased as a rectilinear function of the number of inter-vening packs. The product-moment coefficient of correlation between number of inter-vening packs (x) and average retest reliability (z) was 0.988 and the functional equation relating these two variables was

$$z = 1.886 - 0.239x.$$

Discussion. Under the conditions of this study, esthetic fatigue had the predicted effect upon intra-judge reliability measures. The correlation between a judge's first ranking of a group of esthetic stimuli and his second ranking of the same stimuli is a decreasing rectilinear function of the number of similar groups of esthetic judgments he makes between the two rankings. Whether esthetic fatigue is an example of some other psychological phenomenon, such as retroactive inhibition, is beyond the scope of this study. The study also does not tell us about the kind of change introduced into the judgments by increased esthetic fatigue. About this latter question we assumed that increased esthetic fatigue would increase the random error in judgments, such that the change in judgments from the first to second rankings would be different for each S. It is possible, however, that this change is similar for all Ss with measures of inter-judge agreement being just as high or higher on the second ranking. Since each of the five Ss within our subgroups ranked and re-ranked a different set of stimuli, we cannot test this hypothesis here.

Summary. We have shown that increasing the S's exposure to similar esthetic stimuli results in greater changes in his judgments, but whether these changes are in the direction of greater or lesser conformity to the judgments of other Ss remains unknown.

THE REINFORCING EFFECT OF CHANGES IN ILLUMINATION ON LEVER-PRESSING IN THE MONKEY

By LOUIS E. MOON and THOMAS M. LODAHL, University of Wisconsin

Previous experiments have demonstrated the role of exteroceptively elicited drive in learning and in visual exploration and manipulation.[1] In recent experiments with the rat, the frequency of bar-contacts has been reinforced by increments in level of illumination.[2] In the present experiment, comparable results were sought at the level of the monkey and the effect of direction of illumination-change was studied.

Subjects. Thirty-three immature rhesus monkeys with a mean weight of 5.72 lb. were used in this experiment. The animals were randomly assigned to the experimental and control conditions, and the mean weights of the two groups did not differ significantly.

Apparatus. The apparatus consisted of a box 91 cm. long, 59 cm. wide, and 60 cm. high, constructed of ⅜-in. plywood, and completely lined with 24-gauge sheet-iron painted a dull gray. This box was mounted on a wooden base covered by a metal grid which served as the floor of the box. A distance of approximately 10 cm. separated this grid from the floor of the room. The box was so constructed that no external light could penetrate into the apparatus, while a perforated baffle on the base permitted the free passage of air. An entrance-door was placed at one end of the box, and from the opposite wall a metal lever 8.5 cm. long projected into the box at a height of 34 cm. from the grid-floor and at a positive angle of 40° from the horizontal. To the lever were attached a microswitch and a return-spring, contained within a box externally mounted on the apparatus, and a downward pressure of 86 gm. was necessary to depress the lever far enough to actuate the switch. An outward projection of the wall above the lever prevented *S* from using the lever as a perch. The ceiling of the apparatus contained a screened aperture which led to a centrifugal blower (Dayton Type 10-180), which sucked air out of the box and thus created continual ventilation through the perforated baffles. This blower was in continuous operation whenever *S* was in the apparatus. A translucent screen of frosted Plexiglas about 520 cm.² in area was located in the exact center of the

[1] D. E. Berlyne, Novelty and curiosity as determinants of exploratory behavior, *Brit. J. Psychol.,* 41, 1950, 68-80; H. F. Harlow, M. K. Harlow, and D. R. Meyer, Learning motivated by a manipulation drive, *J. exp. Psychol.,* 40 1950, 228-234; J. P. Flynn and E. A. Jerome, Learning in an automatic multiple-choice box with light as incentive, *J. comp. physiol., Psychol.,* 45, 1952, 336-340; R. A. Butler, Discrimination learning by rhesus monkeys to visual-exploration motivation, *J. comp. physiol. Psychol.,* 46, 1953, 95-98; Butler and Harlow, Persistence of visual exploration in monkeys, *J. comp. physiol. Psychol.,* 47, 1954, 258-263.
[2] J. B. Girdner, An experimental analysis of the behavioral effects of perceptual consequence unrelated to organic drive states, *Amer. Psychologist,* 8, 1953, 354-355; M. H. Marx, R. L. Henderson, and C. L. Roberts, Positive reinforcement of the bar-pressing response by a light stimulus following dark operant pretests with no after effect, *J. comp. physiol. Psychol.,* 48, 1955, 73-76.

ceiling and protected on the inside by wire mesh. Two Westinghouse lamps, 60 and 15 w., respectively, were mounted on the outside of the apparatus behind this translucent screen and served to illuminate the box. A one-way observation window, 7 cm. square, was placed between the illumination-screen and the side of the apparatus which contained the entrance-door. Closure of the microswitch by depression of the lever effected any stimulus-change required by the experimental conditions and simultaneously activated an electric counter (Veeder-Root Type B-120506) located on the control-panel in an adjacent room.

Experimental design. To control for the effect of direction of stimulus-change, it was necessary to divide the Ss into four balanced groups. The 20 Ss in the experimental group were randomly assigned to two subgroups, each containing 10 animals. One subgroup was subjected to condition 'increment,' in which the 15-w. lamp remained on when the lever was not depressed and depression of the lever substituted the 60-w. lamp for the 15-w. lamp, thereby markedly increasing the level of illumination inside the box. The other subgroup experienced condition 'decrement,' in which the 60-w lamp was on when the lever was not depressed and depression of the lever substituted the 15-w. lamp, thus effecting a sizable decrease in the level of illumination.

The 11 Ss in the control group were randomly assigned to two subgroups, the first (A) containing 7 and the second (B) containing 4 animals. In Condition A, the box was illuminated continuously by the 60-w. lamp; in Condition B, the box was continuously illuminated by the 15-w. lamp. Depression of the lever produced no change of illumination under either condition.

All experimental sessions occurred 15 to 23 hr. after the last feeding the Ss had received. The apparatus was placed in a dark room containing a loud-speaker which emitted thermal noise at a level sufficient to mask disturbing auditory stimuli emanating from other parts of the laboratory. Each S was placed alone in the box, the apparatus was activated from the control-panel in an adjoining room, and the time from the activation of the apparatus to the time of the first lever-depression was recorded. Animals not responding within 10 min. after being placed in the box were discarded, and two monkeys were eliminated on this basis. Each S remained in the apparatus for exactly 60 min. following its first response. At the end of this time the apparatus was turned off, the animal removed, and the total number of responses recorded. Each S was tested only once in the apparatus.

Results. A frequency-plot of all scores yielded a positively skewed distribution which was normalized by a logarithmic transformation. The results are presented in Table I. The combined experimental and control groups differed significantly, but the differences between the two experimental subgroups and the two control subgroups did not approach statistical significance. No significant difference was found between the combined experimental and control groups with respect to the latency of the first response in the apparatus.

These results indicate that there is a significant increase in the frequency of lever-depression by monkeys when each response is accompanied by a change in the level of environmental illumination. The effective factor seems to be change *per se* rather than direction of change. It is difficult to see how these results could be attributed to previous conditioning, secondary reinforcement, or anxiety-drive reduction. The Ss had never before encountered the lever, nor were they ever fed

APPARATUS

AN AUTOMATIC DEVICE FOR THE INVESTIGATION OF OPERANT BEHAVIOR IN FISH

By WILLIAM A. DETERLINE, University of Pittsburgh

In the study of the various phenomena of learning, the conditioning procedure involving a 'free operant' is used extensively. The rate of response seems to be a satisfactory measure for use with any organism for which the experimenter can select a suitable operant and an adequate reinforcing stimulus.[1]

The method, however, has been given only limited application in studies of learning or discrimination in fish. Since fish are members of one of the lowest groups of the *subphylum Vertebrata* a wealth of material may lie here for the comparative or physiological psychologist, particularly for studies in which the simpler structure of the organism might be of importance.

The apparatus described by Haralson and Bitterman appears to be the only application of this technique to the study of behavior in fish, and their procedure is too cumbersome for extended experimentation inasmuch as the reinforcements must be introduced manually by the experimenter.[2] The present study, then, had as its primary objective the design of an automatic apparatus which would be suitable for the study of operant learning and operant discrimination in fish.

There are many species of tropical fish which are easily obtained at stores selling aquarium fish, but none of those used with this apparatus has been found to be as satisfactory as the African 'mouthbreeder,' *Tilapia macrocephala,* the same species used by Haralson and Bitterman in the study cited above.[3] *Tilapia* is a gluttonous eater that does not satiate rapidly

on the food used here as reinforcement, adapts quickly to the experimental situation and develops a consistent, high rate of responding.

The apparatus. Each fish during experimentation is kept in an individual 5-gal. tank measuring 14 in. $\times$ 8 in. $\times$ 10 in. The tank is filled to within an inch of the top and contains no sand, plants, or other ornamentation. A round glass ring, 2 in. in diameter, which floats on the surface of the water, is fastened to the center of one end of the tank by a small suction cup. All food, which is of the dry powdered variety, is presented in this ring that prevents the food from spreading over the entire surface of the tank.

The response selected for conditioning in this apparatus is that of the fish striking an object with its nose or lips. The target is a small wedge of plastic fastened to the end of a rod that projects down into the tank from above to a point approximately 1 in. below the surface of the water in a corner of the tank 2 in. from the glass feeding ring.

The target-rod is suspended through an opening in a wooden platform measuring 15 in. $\times$ 9 in. that completely covers the top of the tank. The platform is supported at both ends and at the back by sheets of plywood, all three of which are 10 in. high, the same height as the tank. The tank itself rests on a sheet of plywood which measures 15 in. $\times$ 9 in., the same as the top platform, hence, when the conditioning box is in place, the three upright sides fit snugly around this base.[4]

Above the opening in the top platform the target-rod is turned at a right angle with a pivot at the point of the angle. The end of the horizontal section of the target-lever rests on the upper of two silver contact leaves in such a manner that any slight pressure against the target depresses the upper leaf to make a contact. This activates a solenoid that operates the food-magazine, dropping a few small grains of powdered food through a hole in the platform. The food drops onto the surface of the water within the glass feeding ring and floats there until eaten. Fig. 1 shows the top platform with the target-rod and, in the lower drawings, the food-magazine.

Along the underside of the top platform are four rows of five 6-v. light bulbs connected in series to a 110-v. source. In addition there is one 110-v., 7.5-w. bulb, wired independently of the other bulbs, at the back end of the platform near the target-rod. The smaller bulbs were kept on continuously except during the discrimitive training, to be described later.

The solenoid of the food-magazine can be activated independently by E by means of a button on an extension cord. This is used during successive approximation-training described below.

Procedure. The fish is put into its experimental tank at least one week prior to the first conditioning trial, and it receives all food in the feeding ring as part of the preliminary habituation.

Each fish is food-deprived for 24 hr. prior to each session of conditioning. The conditioning box is lowered in place over the tank, the small 6-v. bulbs turned on, and the power-source for the solenoid connected. Successive approximation-

---

[4] The top platform plus the three upright sides will be referred to as the conditioning box.

training is then begun, during which the fish is reinforced by E for making progressively closer approaches to the target.

Since the conditioning box rests directly on the tank, the vibratory stimuli furnished by the solenoid each time it is activated are carried directly into the water, where they are sufficiently strong to be detected by the fish. Initially all fish show a
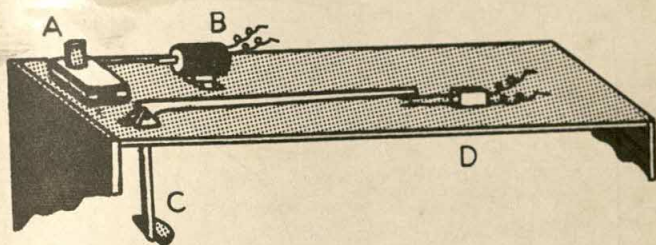
A. Food magazine; B, Solenoid; C, Target; D, Contact points.
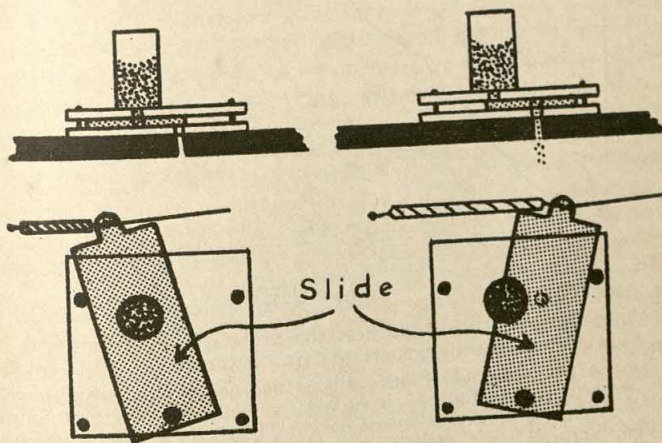
FIG. 1. DIAGRAM OF TARGET-ROD AND FOOD-MAGAZINE

definite startle response to this stimulus, but within a short time this response disappears. The fish learns very quickly to respond to this discriminative stimulus marking reinforcement by moving directly to the feeding ring and taking the food that has been presented. The fish will make successively closer approaches to the target until finally it will touch the target. This will, of course, automatically activate the solenoid and reinforce the fish. From this point on all responses are recorded by E, or by some other form of automatic recorder. When using an electrical recording device, leads are simply connected to the contacts operated by the target-rod. Fig. 2 shows a typical cumulative record made by such an automatic recorder. The slope of the curve indicates the rate of responding.

In this sample record the fish was reinforced on a continuous schedule for 22 min. following the first press of the target. It was then again food-deprived for 24 hr. The second conditioning session consisted of 13 additional minutes of continuous reinforcement after which the fish was put on extinction, *i.e.* the solenoid was disconnected so that no reinforcement was given. After the fish met the criterion of extinction of three consecutive min. without a response the conditioning box was
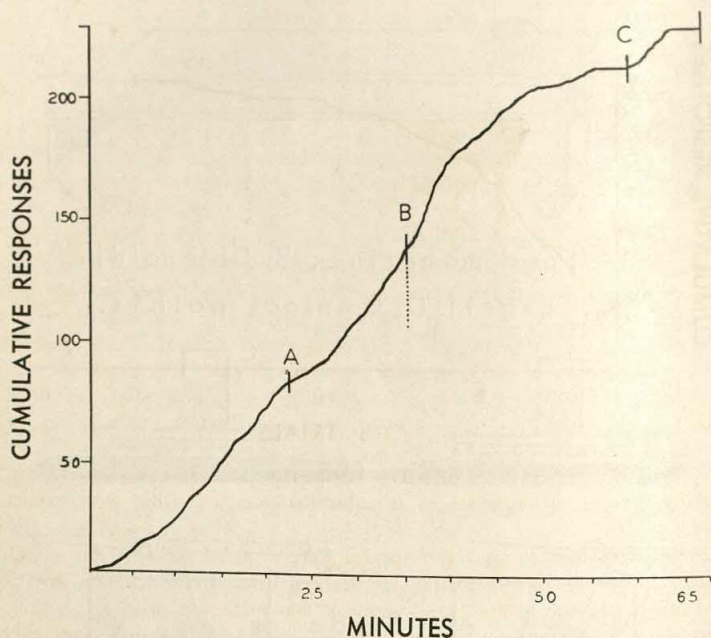
Fig. 2. Acquisition, Extinction and Spontaneous Recovery, Based on Performance of a Single Fish

Vertical line A indicates end of first daily session and beginning of second day. Both days were under conditions of continuous reinforcement. Vertical dotted line B indicates beginning of extinction-session immediately following second acquisition session. Vertical line C indicates point at which criterion of extinction was reached. Portion of the curve to the right of line C represents spontaneous recovery after two hours.

removed. Two hours later the box was again put in place and a measure of spontaneous recovery taken, again to a criterion of extinction of three consecutive min. without a response. This curve is based on the performance of one fish but it is typical of those obtained for other fish in this apparatus.

Operant discrimination. The procedure in a simple discrimination problem was as follows; the light furnished by the small 6-v. bulbs was the positive discriminative stimulus (S^+) in the presence of which all responses were reinforced, and the light from the 7.5-w. bulb, which was very dim in comparison, was the negative

discriminative stimulus (S⁻) in the presence of which no responses were reinforced.

After the initial training described above, the fish was again food-deprived for 24 hr. and the box was placed in position over the tank. S⁺ was turned on and remained on until the fish pressed the target and received food. As soon as the response was made S⁺ was turned off and S⁻ was turned on and remained on for a
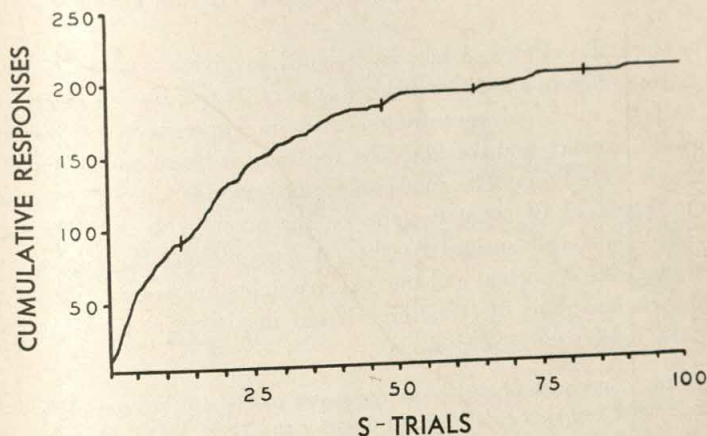
FIG. 3. FORMATION OF A DISCRIMINATION BY A SINGLE FISH
Cumulative record of responses to stimulus that was associated with extinction.

full 2-min. Responses during this S⁻-period were recorded but were not reinforced, the solenoid of the food-magazine having been disconnected as soon as S⁺ was turned off. At the end of the 2-min. period S⁺ was turned on again.

This procedure of alternating conditions was continued, and two measures were taken: number of responses under S⁻, and latency of the response under S⁺. During the early part of the discriminative process the latency[5] was extremely variable, but as the discrimination became progressively more perfect, the latency steadily decreased and became less variable until, by the end of the fourth day, the latency was a very consistent 4 sec.

The curve shown in Fig. 3 represents the cumulative number of responses during successive S⁻ periods during the five daily sessions. The curve indicates a fairly orderly decrease in rate of responding during S⁻ with an almost perfect discrimination during the fifth session, in which only two responses were made during 17 successive S⁻-periods.

The apparatus described above proved adequate for use in the investigation of operant behavior with this class of organism. It should provide a means for detailed studies of behavior of suitable species of fish.

---

[5] The latency measure was time in seconds from onset of S⁺ to the first response.

A METHOD FOR CONTINUOUSLY MEASURING THE POSITION OF A RAT IN A RUNWAY

By A. H. BLACK and R. L. SOLOMON, Harvard University

The apparatus described here is designed to give a continuous record of a rat's position in a straight enclosed runway. It provides an inexpensive and accurate means of measuring activity in experiments on learning, motivation, conflict, and the like. The apparatus is based on the design of the simple scale beam. The runway is placed on one side of the beam's fulcrum and a set of counterweights on the other, with the runway-end suspended on a coil spring. As the rat moves along the runway, the balance between its weight and the counterweights on the other side of the fulcrum changes, and the resulting vertical displacement of the runway accurately reflects the rat's position.

The main components of the seesaw are shown in Fig. 1. The arms of the beam are made up of two steel rods ½ in. in diameter and 72 in. long (A).[1] These rods are 8 in. apart. They are joined by two crossbars 1 in. in diameter. The steel rods extend through holes in the crossbars. One of the crossbars is 1 in. from the left-end of the steel rods (as illustrated) and the second 28 in. from that end. The second crossbar is mounted on a set of ballbearings which are supported by two posts attached to a baseboard. This crossbar acts as the fulcrum (B).

The enclosed runway is mounted on the end of the rods projecting 44 in. to the right of the fulcrum. The goal-box end of the runway (C) is the one nearer the fulcrum. On the other side of the fulcrum, a third ½-in. steel rod is run from the fulcrum to a point 8 in. beyond the first crossbar, and counterweights are mounted on this rod. These weights are simply cylindrical pieces of iron with ½-in. holes bored in the center to allow them to fit on to the steel rod. Set-screws hold them in the desired position.

A spring (D) runs from the top of the starting-box to the ceiling. Thus, the total balancing setup is really a combination of Type-1 and Type-2 levers. The balancing of the rat and runway on one side of the fulcrum against the counter-weights on the other side gives a Type-1 lever. The balancing of the rat and runway on one side of the fulcrum against the spring pulling upward on the same side of the fulcrum gives the equivalent of a Type-2 lever. One might expect the arms of the beam to swing down and touch the lower stop each time the balance of the system is upset by the movement of the rat, but the use of a spring to balance the runway and rat prevents this complete swing with each movement. A piston (E), attached to the bottom of the runway at the starting-box, runs up and down in a

---

[1] Letters in the text refer to those of Fig. 1.

cylindrical tube containing light mineral oil. It serves to damp oscillations which might otherwise be set up in the suspension-spring. Stops are arranged at both ends of the lever to prevent the vertical displacement of the lower arm from exceeding 4½ in.

The enclosed runway is a cage constructed of hardware cloth 48 in. long, 8 in. wide and 8 in. high. The goal-box section (C) is at the end of the runway nearest
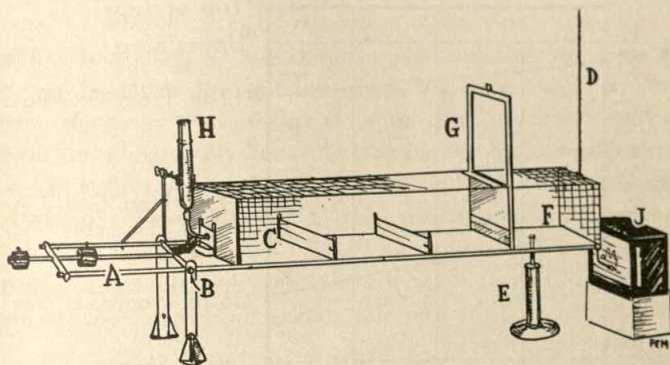
FIG. 1. DRAWING OF THE APPARATUS

the fulcrum and 3½ in. from it; the starting-box (F) is at the other end. The ends of the runway are removable pressed wood panels, 8 in. square. The top of the cage above the goal-box and above the starting-box is hinged in 12-in. sections to allow easy access to the interior of the runway. Since the runway must be free to move vertically, the gate of the starting-box (G) has all its parts attached to the runway. Two grooved slats of wood, used as runners for a thin aluminum door, are attached to the sides of the cage, 12 in. from the starting-box end. The door is suspended by a spring from a light crossbar which connects the two runners. The crossbar is 12½ in. above the top of the runway. When a catch keeping the door down is released by a solenoid, the door is pulled up above the runway by the spring and held there between the two runners for the duration of a trial.

Since the purpose of this apparatus is to measure free running there is no proper goal-box in the sense that the rat is closed up in it when he enters. The rat is free to move up and down the runway for the duration of a trial once the starting-gate is open. This freedom presents the problem of keeping the amount of reinforcement on any given trial constant and of measuring time of the occurrence of reinforcement. To solve this problem, water is used as a reward, and amount, rate, and time of reinforcement are measured with a drinkometer.[2] The water-source is a 100-cc. graduate (H) mounted immediately over the fulcrum. It is attached to the fulcrum and filled from the top, since moving it would upset the balance of the system. A drinking tube is attached to the base of the graduate and leads into the runway through a small hole in the end-panel.

[2] J. H. Hill, and E. Stellar, An electronic drinkometer, *Science,* 114, 1951, 43-44.

Records of activity are made by a stylus, attached to the runway at the starting-end, which marks the vertical displacement of the runway on a wax-paper kymograph (J) placed beside the runway. The problem of locating the position of the rat in the runway from these records is solved as follows:

The apparatus is so balanced that it rests at a horizontal position when the runway is empty, and it swings down and comes to rest, just touching the lower
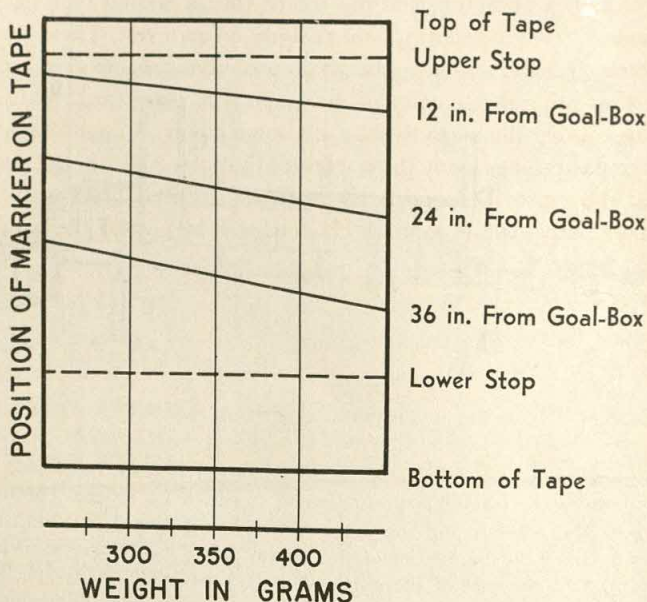
Fig. 2. Calibration of the Apparatus
Method of determining, from the graphic records, the position in the runway of rats of different weights.

stop, when a standard 350-gm. weight is placed at the back of the start-box. (The weight is chosen arbitrarily; any weight, within limits, will do. If the rats used differ greatly in weight, different standard weights and reference-charts should be used for each weight-group.) The standard weight is then moved to different positions in the runway, and records are made on the kymograph of the position of the stylus when the apparatus comes to rest under each condition of weight and position. From these records, it is possible to make graphs showing the position of the rat in the runway when the weight of the rat and the position of the stylus-mark on the kymograph tape are known. An example of such a graph is given in Fig. 2. The abscissa gives the weight of the rat, the ordinate gives the position of the stylus-mark on the tape, and the parameter indicates the position of the rat in the runway. For example, if the stylus-mark is 3 in. above the bottom of the tape, and the rat weighs 300 gm., then the rat is 36 in. from the goal-box end of the runway.

Several cautions, however, are necessary to insure a high degree of

accuracy of these graphs. First of all, two graphs are essential, one for movement from the goal to the start, and one for movement from start to goal since the friction in the system makes the runway stop in slightly different positions in each case. The record for the first type of movement can be made by releasing the runway from the goal-box position on each weight- and position-test, and the record for the second type by releasing the runway from the starting-box position on each test. These precautions are necessary, however, only when high precision is required. A useful technique for testing the accuracy of the graph is to place small barriers in the runway, causing the rat to hesitate at known points. A comparison of these kymograph-readings with those expected on the basis of the graphs described above provides a fairly adequate test of the accuracy of the graphs. For most purposes, the accuracy obtained will be beyond the requirements of E.

A MICROPOLYGRAPH

By VLADIMIR CERVIN, University of Toronto, and BOHDAN CERVINKA,
Lausanne, Switzerland

This paper describes the design and function of a small, portable,
multiple-recording apparatus which we call a 'micropolygraph,' and which
was designed to overcome the limitations of existing polygraphs, especially
their large size. The fact that time is the basis of measurement and com-
parison of observations, suggested to us that a combination of a wrist-
watch and multiple-recording device would give the required type of
apparatus. Such an apparatus has now been designed for recording obser-
vations of bodily movements, pulse-rate, and galvanic skin response
(*GSR*), as well as local time. It can be adapted to record other mechanical,
acoustic, and electrical signals as well.

The following paragraphs describe the general characteristics of a micropolygraph
with two transducing devices—a mechanical one for bodily movements, and an
electromagnetic one for *GSR* and pulse-rate. The micropolygraph as designed con-
sists of an ordinary movement of a wristwatch, and two transducers linked to two
recording tools, *e.g.* styli or needles, registering the observations on a moving tape
which is synchronized with the watch-movement during recordings.

The design of the device for recording bodily movements, shown schematically in
Fig. 1, provides for a freely rotating weight similar to those used in self-winding
watches (with no angular-movement limits), mounted in the center of the watch
under the dial, and bearing an eccentric with a follower attached to a pushrod, which
is mechanically linked with a countershaft bearing a stylus. The stylus touches
the sensitive side of the recording tape, which is coated with a substance suitable
for receiving and preserving the recording made on it. The tape, calibrated in
intervals of 1 sec., is wound on two reels, one supply (or driven) reel, and one
take-up (or driving) reel, driven by its own spring. A release is provided for
starting the reels, by coupling the take-up reel to the escapement-mechanism of the
watch. When no recordings are desired, the take-up reel is put out of gear and
locked by the same device. As the weight oscillates, the countershaft rotates to and
fro, moving the stylus up and down. The stylus thus traces a curve on the tape
whenever the tape is moving.

In Fig. 1, eccentricity is exaggerated for clarity; the multiplication-factor of
linkage between the eccentric and the stylus is distorted for the same reason; the
spring-loading of the eccentric follower is omitted for simplicity. The scale of the
two views is somewhat distorted: the vertical scale is approximately 2.5 : 1, while
the horizontal scale is 2 : 1.

With the present size of watches, reels carrying up to 15 ft. of tape, 4 mm.

wide, can be used. The speed-limits of the tape, which depend on the diameters of the reel when empty and full, will be between 0.73 mm./sec. and 0.42 mm./sec. With these specifications, the total recording time of one reel will be from 2-3 hr. As the graphs obtained will be small, special reading and reducing devices will have to be used.

The recordings described will show the duration, intensity, and perhaps the
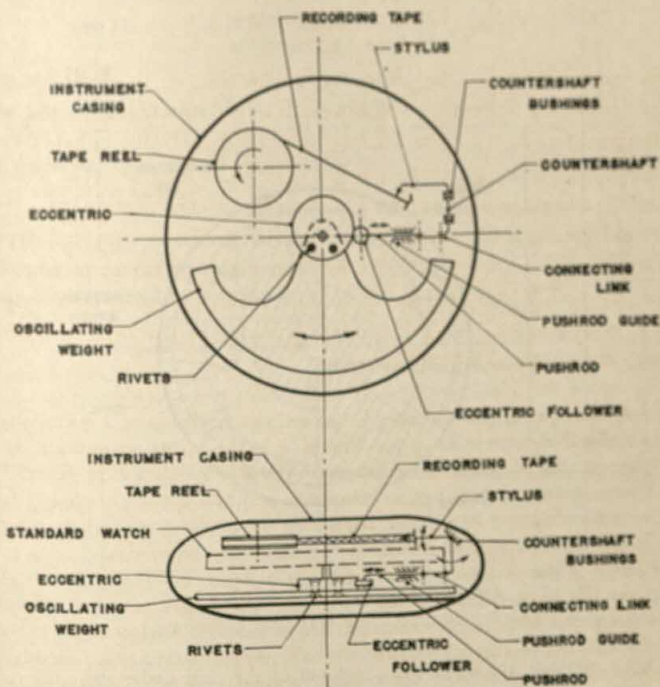
FIG. 1. PLAN AND ELEVATION OF THE MICROPOLYGRAPH
(Device for recording bodily movements.)

direction of the movements of the bodily part on which the apparatus is worn. Since the micropolygraph can be worn on any part of the body, it may replace, or complement to a certain extent, the tremograph, ergograph, automatograph, and ataximeter.

Recordings of *GSR* and pulse-rate can be combined by recording the pulse as short 'blips' on the *GSR*-curve. The design of the transducer is schematically shown in Fig. 2. Direct current, passing from a 1-v. cell through the skin and a cardio-microphone, is fed into two hollow coils. Two iron cores fixed on a rotating bridge are designed to penetrate into the coils. The bridge is provided with an actuating pin, mechanically linked to a second stylus in such a way that when the cores are fully in, the stylus is at its highest point on the tape and when they are out, at its

lowest point. As the current in the coils fluctuates, the cores move in and out, rotating the bridge and moving the stylus, which traces a curve on the tape when the latter is moving. The transducer is mounted in the back lid of the watch.

The circuit will be so arranged that with maximal current passing through the coils, *i.e.* when skin-resistance is minimal and conductance maximal the cores will penetrate into the coils to a maximal length. When skin-resistance is maximal, and
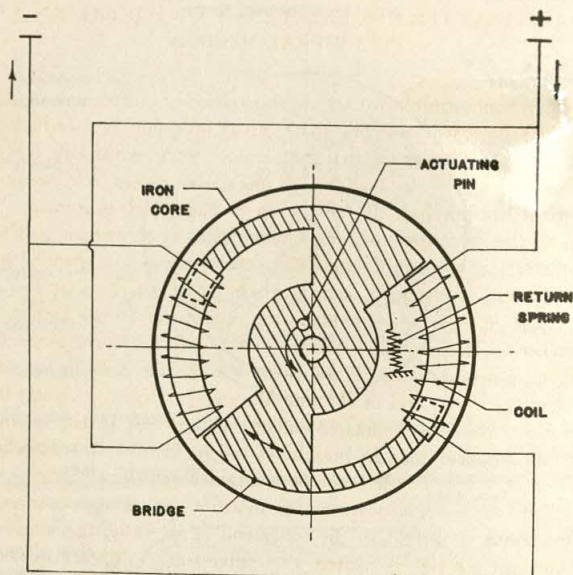
Fig. 2. PLAN OF THE TRANSDUCER
(Device for recording *GSR* and pulse-rate.)

conductance minimal, the cores will move out of the coils under the action of the return-spring. Adjustment of the recording range can be provided, if necessary. The device therefore will record the absolute level of skin-conductance.

Unless the size of the watch is unduly increased, it seems that only the above three recordings can be made simultaneously. Interchangeable lids with additional transducer-elements may, however, be provided for other recordings.

The micropolygraph, when manufactured, will make possible quantitative observations of human and animal behavior under conditions where available apparatus cannot be used, for example in everyday life or in miniature laboratory situations. Moreover, even in experiments where present apparatus can be used, the micropolygraph offers the possibility of considerable improvement of the experimental control of the irrelevant variable which is the apparatus itself. The micropolygraph will minimize the stimulation derived from the recording apparatus.

APPARATUS NOTE

A SIMPLE METHOD OF DEMONSTRATING DIFFERENCES IN THE CRITICAL FLICKER FREQUENCY OF FOVEAL AND PERIPHERAL VISION

The only equipment required for the demonstration of differences in critical flicker frequency (*CFF*) of foveal and peripheral areas of vision is a cathode ray oscilloscope that is equipped with an internally generated sweep of variable frequency. Since an oscilloscope of this type is available in most laboratories, the demonstration can be made without the purchase or fabrication of additional equipment.

In setting up the demonstration, the sweep-circuit is turned on and the sweep is centered on the face of the oscilloscope. Then the sweep is so expanded that the ends are not visible; this to eliminate certain physical irregularities at the ends of the sweep. The result of these adjustments is the production on the face of the oscilloscope of a flickering line, in which the frequency of flicker is the same at all points. The sweep-frequency should be so adjusted that the physical flicker can be seen. Then while looking at one end of the line, the sweep frequency should be increased. When visual flicker ceases at the fixated end of the line, there is still an obvious flicker at the other end which is seen in peripheral vision. In other words, when *CFF* has been reached for the fovea it has not been attained for the periphery.

The possibility that the results are due to an artifact in the apparatus can readily be excluded. If fixation is shifted to the other end of the line, the effect is reversed, and if the center of the line is fixated, the center section appears steady and flicker is seen at both ends. By simply turning the head at various angles, it can also be demonstrated that the same principle holds for all meridians of the visual field.

Since the *CFF* varies somewhat among *O*s, best results are obtained by having each *O* adjust the sweep-frequency for his own observations. This procedure has the additional advantage of placing *O* so close to the oscilloscope that the face of the oscilloscope subtends a large visual angle.

KENNETH T. BROWN

Brown University

NOTES AND DISCUSSIONS

ON THE NUMBER OF ARTICLES OF PSYCHOLOGICAL INTEREST PUBLISHED IN THE DIFFERENT LANGUAGES: 1946-1955

Four times before, in 1917, 1926, 1936, and 1946, the writer has published analyses of the number of articles published in the different languages.[1]

The first article covered the period of 1894-1915; the second was a supplementary report covering the 10-yr. period 1916-1925; the third and fourth were additional supplementary reports covering the periods of 1926–1935 and 1936–1945. The present paper is an additional report analyzing the psychological titles for the 10-yr. period just closed; namely 1946–1955. In the first three reports, the material consisted of analyses by language of the titles listed in the *Psychological Index,* which was discontinued in 1936. The materials for the fourth and the present reports were obtained from the *Psychological Abstracts.*

The 1917 report may be considered to cover the development of psychological publication from the relatively early days to World War I. The 1926 report covers the period of this war and the post-war years. The third report of 1935 showed the trends of psychological publication in the mid-war period of readjustment and of the economic depression which grew out of the readjustment. The next report, of 1945, covers the years of growing Fascism in Europe and of the years of World War II. The present report of 1955 covers the periods from the end of the Second World War, through the difficulties in Korea and marks, we believe, a period of increased feeling of nationalism in many countries.

One source of error, which is eventualy self-corrective, is that the titles are listed under the year of publication in the *Abstracts* rather than the actual year of publication of the book or article. There is sometimes, but not frequently, a lag in listing of one year, but the error caused by this lag is so small as certainly not seriously to effect trends. Another source of error may have to do with reports of a classified nature having to do with war potential in various countries.

---

[1] S. W. Fernberger, this JOURNAL, 28, 1917, 141-150; 37, 1926, 578-580; 48, 1936, 680-684; 59, 1946, 284-290.

The numbers of articles in English, German, French, Italian, Russian, and 'All others' for each of the 10-yr. periods are given in Table I. The results for every year are found in successive rows and for each language in successive columns. In the last column will be found the total number of titles in all languages. The variation in the total number of titles may be not wholly dependent upon the number of books and articles which were listed during any given year. The total is partially dependent upon the

TABLE I

NUMBER OF TITLES IN DIFFERENT LANGUAGES FOR EVERY YEAR DURING 1946–1955

Year	English	German	French	Italian	Russian	All Others	Total
1946	3930	235	207	97	167	300	4936
1947	4254	171	105	5	8	125	4668
1948	4817	206	233	54	30	272	5612
1949	5425	296	353	32	98	326	6530
1950	5735	108	333	35	114	238	6563
1951	7017	318	419	85	168	315	8322
1952	5673	379	579	98	279	289	7297
1953	6423	558	530	56	156	360	8083
1954	7795	469	429	85	106	236	9120
1955	7554	572	351	61	82	483	9103
average	5862.3	331.2	353.9	60.8	120.8	294.4	7023.4
high	7795	572	579	98	279	483	9120
high year	1954	1955	1952	1952	1952	1955	1954

inclusion or elimination of borderline fields, as the result of gradually changing editorial policy.

One immediately interesting fact is the enormous increase in the total number of titles. This increase is from the low year of 1947, with only 4668 titles, to the high year of 1954 with 9120 titles—an increase of over 95%. From a study of the articles themselves, it would seem that two factors are the basis for this increase: (1) better coverage by the *Abstracts;* and (2) an enormous increase in the number of psychological journals and the number of articles appearing in those already established before the beginning of this period. The year 1955 maintains the high level reached the year before.

The results from the English studies show a similar marked increase from 3930 titles in 1946 to the high of 7795 in 1954—an increase of more than 93%; a level maintained in 1955. The increase from year to year is steady, except for an inversion in 1951 when the number of titles was greater than any of the years before or for the two years following. This may have been based on editorial policy of collecting titles of previous years, which had not been available, or it may have been due to an enormous increase in volume of publication after World War II.

The results for titles in the German language show an enormous difference in number and a somewhat different distribution than those in English. The lowest number of German titles occurs in 1950; namely 108. Since that year there has been an almost constant increase to 572 in 1955. This final number is greater than for any year since 1940.

The distributions of titles in French are, for 5 of the years, greater than for those in German. The low is in 1947 with only 105 items, and the high in 1952 with 579 titles. Except for the low year of 1947, the number of French titles is greater each year than the number for the last 5 years of the previous decade (1941–1945), the period of World War II and the German occupation.

The number of titles in Italian has dropped markedly since 1941 and, although there has been a slight recovery, in no year was the 1940 level reached during the last 10 years. For this period, the low is 5 titles in 1947, and the high 98 titles in 1952.

The number of titles in Russian has been very uneven during this 10-year period. The low was 8 titles in 1947 and the high 279 in 1952. Since that date there has been a steady decrease in Russian titles, with only 82 in the last year (1955).

The number of titles in languages other than those listed above show a relative but not steady increase during the decade, 1946–1955. The low of 125 items occured in 1947 and the high in 1955 with 483 titles. There was only one year in the previous decade (1936) with more titles than 1955.

In previous publications the writer has attempted to show the effects of politics, economics and war on the volume of scientific publications in psychology.[2] When one considers the post-war history of Germany, France, and Italy, as compared with that of the United States, it would seem that one can readily account for the differences and trends in the volumes of scientific publication during the last 10 yr.

It is of interest to analyze the distribution of languages grouped together under the category 'All others.' The number of titles is given below, and the number in parentheses indicates the number of years in which there was publication listed in that language. In all, 18 languages are represented as follows for the total 10-yr. period: Spanish 977 (10); Scandinavian (Swedish, Norwegian and Danish) 607 (10); Dutch 381 (10); Portuguese 234 (10); Hebrew 197 (7); Polish 139 (8); Japanese 110 (6); Arabic 97 (4); Czech 43 (5); Finnish 35 (7); Boer 16 (3); Turkish 11 (2); Hungarian 10 (5); Greek 10 (2); Roumanian 8 (3); Chinese 3 (1); Serbian 3 (2); and Bulgarian 1 (1).

---

[2] Fernberger, Publications, politics and economics, *Psychol. Bull.*, 55, 1938, 84-90; Scientific publication as affected by war and politics, *Science*, 104, 1946, 175-177.

Publications in Japanese and in Polish have been noted only during the last 6-8 yrs. Whether this is due to better productivity in these countries, or to better coverage by the *Abstracts* cannot be determined. A similar increase in recent years is true of publications in both Hebrew and Arabic. There seems to have been a marked falling off of publication among the Russian Satellites, as compared with the previous decade, especially in

TABLE II

BREAKDOWN OF ENGLISH LANGUAGE TITLES BY COUNTRY OF AUTHOR

Year	Total English	Not Identified	American	Total British	British Isles	British Dominions and Colonies	Other Countries
1946	3930	16	3469	397	295	102	48
1947	4254	19	3751	443	289	154	41
1948	4817	20	4288	468	326	142	41
1949	5446	20	4780	552	409	143	94
1950	5752	17	4998	593	412	181	144
1951	7034	17	6059	763	618	145	195
1952	5692	19	4896	599	451	148	178
1953	6435	12	5466	777	541	236	180
1954	7795	20	6724	828	602	226	223
1955	7572	19	6591	842	571	271	120

Hungarian (from 115 titles to 10 during the last two 10-yr. periods). The exception among the satellites is an increase in Polish with 172 titles as compared with 139 in the previous decade. It should be noted that many of the titles in Spanish and Portugese originate in South America—especially in the Argentine and Brazil. It should also be noted that only publications in Spanish, Scandanavian, Dutch, and Portugese have been included in each of the years in the decade.

One additional analysis seems of interest; namely a breakdown by countries of the authors of the English language articles. One identifies the author by recognition of the name, by the place of publication (and more recently much simpler inasmuch as the editors of the *Abstracts* have added the place of occupation, when known), and by research by the writer. Still a number of individuals failed to be identified by country. The results of this analysis will be found in Table II. The largest number of English titles originated in the United States and this number has been increased almost steadily from 3469 in 1946 to a high of 6724 in 1954. There are reversions in 1952 and 1953 which, it should be pointed out, is the period of the Korean 'Police Action.' The number of British titles is quite small, for any year, and varies from a low of 397 items in 1946 to a high of 842 in 1955. The increase has been fairly but not completely steady. This increase in the number of titles of British origin is to be accounted for

largely in terms of publications within the British Isles. Here the low was in 1947 with 289 items and a high of 618 in 1951, a high which has remained fairly constant during the subsequent four years. The contributions from the British Dominions and Colonies shows a relatively steady increase from 102 items in 1946 to a high of 271 in 1955. Most of these originate in Canada, India, Australia, and New Zealand. The titles in English, found in the last column of Table II were produced by nationals

TABLE III

PERCENTAGE OF TITLES IN DIFFERENT LANGUAGES

Year	English	German	French	Italian	Russian	All Others
1946	79.6	4.8	4.2	2.0	3.4	6.1
1947	91.1	3.7	2.3	0.1	0.2	2.7
1948	85.8	3.7	4.1	1.0	0.5	4.9
1949	83.2	4.5	5.4	0.5	1.5	5.0
1950	87.4	1.6	5.1	0.5	1.7	3.6
1951	84.3	3.8	5.0	1.0	2.0	3.8
1952	77.7	5.2	7.9	1.3	3.8	4.0
1953	79.5	6.9	6.6	0.7	1.9	4.4
1954	85.5	5.1	4.7	0.9	1.2	2.6
1955	83.1	6.3	3.9	0.7	0.9	5.3

of countries other than Britian and the United States. The number of such titles has increased, over the decade, from a low of 41 items in 1947 and 1948 to 223 items in 1954. During the latter years of the decade much of these publications in English have come from Scandanavia and Japan.

For the sake of completeness, a third table is added in which will be found the percentages of titles in the different languages. Obviously, the same trends will be found here as were previously discussed for the data in Table I. Obviously there is a great superiority for English language publication. This never drops below 77.7% of the total (in 1952) and rises as high as 91.1% in 1947. One can say that, during the decade, more than 80% of all publications were in English and, of them, a very large percentage were produced by psychologists within the United States. Except for a few reversions the percentage of titles in German, French, and Italian have varied relatively slightly over the decade and this is equally true for titles under the category All others.

When the present writer made the first analysis of these matters in 1917, he attempted to impress upon the graduate student in psychology, the necessity of a reading knowledge of English, German, and French, no matter what his native language. In the very early times, there were years in which there were actually more publications in German and

in French than in English. Certainly important articles are still published in languages other than English, but the English has gained such a numerical superiority that the reading knowledge of other languages is less important than it was some decades ago. The tables show that psychologists today need not be competent in the 23 languages listed above, but that he must have an adequate reading knowledge of English to be able to keep abreast of advances in his science.

University of Pennsylvania   S. W. FERNBERGER

ERRATA OF FREUD

In this Note I examine two of Freud's philologisms and point out how dangerous half-knowledge may be in the process of pursuing scholarly aims.

In his *Interpretation of Dreams* and other publications on the nature of sexual desire, in which he tried to establish a direct connection between the symbolic visions of the dreamer and the symbolic actions of the mentally deranged, Freud classified various objects of the dream into several groups and attached to them definite rational leanings of man's conscious behavior. Freud reached this equation between symbols and their meanings by various proofs and arguments, such as external appearance and semantic association of individual words and expression. All such variegated substantiations led Freud to believe that "the dream is the (disguised) fulfillment of a (suppressed, repressed) wish," which in the majority of cases is of the 'libidinal' nature.[1]

Realizing the necessity of establishing a rationally conceivable equation between the symbol and its meaning, Freud also tried to utilize linguistics —but with this he not only failed but also created a considerable doubt as to the validity of his other non-linguistic explanations. While speaking about latent meaning of the wood symbol in a dream, Freud wrote:

Now, in the Atlantic Ocean, there is an island named Madeira, and this name was given to it by the Portuguese when they discovered it, because at that time it was covered with dense forest; for in Portuguese the word for wood is *madeira*. But you cannot fail to notice that this *madeira*, is merely a modified form of the Latin *materia* which again signifies material in general. Now *materia* is derived from *mater* a mother, and the material out of which anything is made may be conceived of as giving birth to it. So, in the symbolic use of wood to represent woman or mother we have a survival of this old idea.[2]

From mother, by means of the Oedipus complex, Freud very easily

[1] *The Basic Writings of Sigmund Freud* (A. A. Brill, ed.), 1937, 235.
[2] Sigmund Freud, *A General Introduction to Psychonalysis,* 1949, 143.

reached the desired conclusion. Let us pause, however, and see whether there is a linguistic connection between *Madeira* and *mater* as Freud wishes us to believe. Spanish or Portuguese *madeira* really derives from the Latin *materia,* but the derivation stops right there because *materia* derives from the Indo-European root *dmā* designating the action of building or construction in general. Hence there are such verbal formation as *dmāteriēs, dmāter* and even *domus.*[3] Freud arrived at his false notion about the historical origin of *materia* either on the basis of analogy or took it from some linguists who also reasoned along the same lines.[4]

Being aware that such linguistic comparison is not applicable to other languages, Freud decided to prove his supposition also with the German language. He wrote:

> There is still something to be said on the subject of wood. It is not easy to see why wood should have come to represent a woman or mother, but here a comparison of different languages may be useful to us. The German word *Holz* is said to be derived from the same root as the Greek *üle,* which means stuff, raw material. This would be an instance of process which is by no means rare, in that a general name for material has come to be applied to a particular material only.[5]

As one can observe, here Freud tried to be cautious by using an expression "is said to derive," indicating that he was not certain about it. This uncertainty, however, did not prevent him from making a definite conclusion, *i.e.* from implying that between wood, as a dream-symbol, and the sexual desire there exists some connection. If Freud had been acquainted, at least in general lines, with the problem of the 'consonant-ismus' of the Indo-European languages, he would never have made such a mistake, knowing that the German consonant *h* (*Holz*) derives from the Indo-European aspirated *k.*[6] The German word *Holz,* which through its various historical gradations (Teutonic type *hultos,* pre-Teutonic *kldos,* etc.) is of the same origin as the Greek *klados* (twig) or the Old-Slavonic *klada.*[7] Thus the word *üle* has nothing in common with the word *Holz,* although it conveys a similar meaning. What linguistic relation does Freud see between *üle* and *mother* is hard to imagine.

Even if Freud's reasoning were correct, if this linguistic connection

---

[3] For a detailed treatment of this problem see Alois Walde, *Lateinisches Etymologisches Wörterbuch,* 1906, 373; also K. Brugmann, *Vergleichende Laut, Stammbildung und Flexionslehre der Indogermanischen Sprachen,* 1897, 357.

[4] See, for example, A. Ernout et A. Meillet, *Dictionnáire Etymologique de la Langue Latine,* 1939, 596.

[5] Freud, *op. cit.,* 142.

[6] See Hans Krahe, *Indogermanische Sprachwissenschaft,* 1948, 68; also Herman Hirt, *Etymologie der Neuhochdeutsche Sprache,* 1921, 13, 191.

[7] Consult Klege's *Etymological German Dictionary,* Trans. by J. F. Davis, 1897, 157.

existed, as he assumed, how would the dreamer, who lacks a training in linguistics, realize it? Freud does not worry about this. He wrote:

I do not maintain that the dreamer knows this; on the other hand, I contend that there is no need for him to know it. He probably knows something else from having been told it as a child, but even this, I will maintain, has contributed nothing to symbol-formation.[8]

Why then, should one have recourse to a field with which one is not familiar if it "contributes nothing" to the explanation of the phenomenon in question? Analogy in linguistics, sometimes, may be very misleading and one really must be thoroughly acquainted with the intrinsic laws of general linguistics to avoid mistakes such as Freud made. A logical question to ask here would be how should one explain the wood-sex equation in languages in which there is no such analogy as exists in German and Spanish? For example, the French word *bois* is of uncertain origin. It derived either from the West Germanic *bosk* (English bush) or the Latin *boscum,* and consequently has nothing to do with the Latin *materia* or *mater.*

Such mistakes as these seriously undermine one's confidence in Freud's attitude toward the study of dreams and force one to believe that it is not, as Tridon wished us to believe, that of a "statistician who does not know, and has no means of foreseeing, what conclusions will be forced on him by the information he is gathering, but who is fully prepared to accept those unavoidable conclusions."[9] It seems to us that Freud had "a preconceived bias, hoping to find evidence which might support his views."[10]

University of Notre Dame   JOHN FIZER

TWENTY-SEVENTH ANNUAL MEETING OF THE EASTERN PSYCHOLOGICAL ASSOCIATION

The Eastern Psychological Association met March 23 and 24 at Chalfonte-Haddon Hall in Atlantic City, New Jersey. A total of 1525 persons registered at the meetings. Of these, 877 were members of the Association, 412 were guests, and 236 were new members who joined the Association at the meetings. The present active membership of the Association totals 2763.

A committee headed by Elliott Danzig was responsible for local arrangements and the program was planned by a committee under the chairman-

[8] Freud, *op. cit.,* 143.
[9] Freud, *Dream Psychology,* Intr. by André Tridon, 1921, v.
[10] *Idem.* v.

ship of Malcolm G. Preston. The program consisted of 164 papers (presented in 21 sessions), 5 symposia, 4 special meetings, 3 invited addresses, and 3 films. The 21 sessions were concerned with the following topics: brain functioning, social psychology, applied psychology, visual perception, verbal learning, general experimental, psychotherapeutic processes, central sensory mechanisms, vision, sensory (non-visual), measurement, animal learning (3 sessions), projective techniques, clinical assessment, comparative, avoidance learning, personality, developmental, and human learning. The symposia were as follows: "The impact of the mental health orientation on the future of clinical psychology"; "Transference and countertransference in psychotherapy"; "Human engineering: Its scope and focus; perspect and prospect"; "The psychologist's role in social agencies"; and "Some determinants of behavior in two-choice experiments." Special meetings included a meeting of Psi Chi, and meetings on "Clinical problems in counseling the handicapped"; "Psychology in advertising and marketing"; and "Aptitudes and abilities of the clinical psychologist in private practice." The following invited addresses were presented: "The present situation in brain physiology" by Wolfgang Köhler; "The mechanisms of adaptive behavior" by Alfred L. Baldwin; and "Recent studies of effects of brain injury in man" by Hans-Lukas Teuber. Films were as follows: "Psychogalvanometric response in ducklings"; "Visual perception in animals"; and "Margin of safety: Psychological distance under danger."

Dr. Clarence H. Graham presented the annual presidential address, "Some relations of sensation and perception." During the business meeting, it was announced that the following new officers had been elected: President: Fred S. Keller; Board of Directors; George A. Miller and Neal E. Miller; and Treasurer: Roy Hackman.

The 1957 meetings of the Association will be held at the Hotel Statler in New York City April 12 and 13. The 1958 meetings will be held in Philadelphia.

University of Delaware   GORHAM LANE

FIFTY-SECOND ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The fifty-second annual meeting of the Society of Experimental Psychologists was held at Allerton House, University of Illinois on April 2 and 3, 1956. Lyle Lanier, Chairman of the Society for the year, presided at the business meeting and at the sessions for presentation of scientific papers.

The following members attended the meetings: Bartley, Bray, Brogden, Buxton, Dallenbach, Fitts, Garner, Geldard, Grant, Harlow, Helson, Hunt, Kappauf, Kennedy, Landis, Lanier, Licklider, Marquis, Melton, Nafe, Neff, Newman, Pfaffman, Shepard, K. U. Smith, Solomon, Spence, Underwood, Volkmann, and Woodrow.

A. Chapanis, L. G. Humphreys, and D. D. Wickens were elected to membership in the Society. The total membership now includes 57 members and 22 fellows.

Scientific sessions were held during the afternoon and evening of April 2 and morning and afternoon of April 3. Members presented informal reports of research in progress.

The Warren Medal for 1956 was presented to Harry F. Harlow "for a series of brilliantly conceived experiments on the behavior of monkeys, including studies of motivation, learning and problem solving."

The Society accepted the invitation of Harvard University to hold the 1957 meeting in Cambridge. E. B. Newman was elected Chairman for 1956-57.

University of Chicago

W. D. NEFF

FORTY-EIGHTH ANNUAL MEETING OF THE SOUTHERN SOCIETY FOR PHILOSOPHY AND PSYCHOLOGY

One hundred and sixty-two members registered for the forty-eighth annual meeting of the Southern Society for Philosophy and Psychology held in Asheville, North Carolina, April 29, 30, and 31. The Duke University Departments of Philosophy and Psychology and the University of North Carolina Philosophy Department were hosts.

The main address was given by Kenneth W. Spence, Professor of Psychology, University of Iowa, on "The empirical basis and theoretical structure of scientific psychology." Herbert Hochberg, Professor of Philosophy, Northwestern University, replied to Dr. Spence's paper. The presidential address was given by Marion E. Bunch, Washington University, on "The concept of motivation."

The Society elected the following psychologists to full membership: H. D. Baker; W. B. Bierbaum; J. L. Chambers; C. L. Darby; G. W. Meier; E. L. H. Newbury; L. J. Peacock; E. W. Straus, and K. S. Wagoner.

The following psychologists were elected to associate membership: C. C. Durkee; R. L. Frye Jr.; M. A. McCormick; D. F. Pennington Jr.; E. R. Ray; B. B. Ridge, and A. H. Smith.

The new officers elected were: President, William S. Weedon, Univer-

sity of Virginia; Secretary, Wilse B. Webb, U. S. Naval School of Aviation; Treasurer, Sam C. Webb, Emory University.

The psychologist elected to the Council was Rolland Waters, University of Florida.

The 1957 meeting will be held at Gatlinburg, Tennessee.

Georgia Institute of Technology   JOSEPH E. MOORE

SECOND ANNUAL MEETING OF THE SOUTHEASTERN PSYCHOLOGICAL ASSOCIATION

The Southeastern Psychological Association held its second annual meeting April 29-May 1, at the Biltmore Hotel, Atlanta, Georgia. Official attendance figures were 325. The program was made up of seven symposia, 43 papers, films, and the presidential address entitled "Selecting and connecting" presented by John B. Wolfe. Officers for 1956-57 are: President, Nicholas Hobbs; President-Elect, Edward E. Cureton; Secretary-Treasurer, M. C. Langhorne; Members-at-large on the Executive Committee are Dorothy C. Adkins, C. H. Calhoon, J. F. Dashiell. The 1957 meeting will be held at Nashville, Tennessee, with Fisk University, George Peabody College, and Vanderbilt University as sponsors. The 1958 meeting will be held in Atlanta, Georgia.

Emory University   M. C. LANGHORNE

AN ACKNOWLEDGMENT

The JOURNAL is indebted to Mrs. Margaret Russell Bentley for the photograph and signature of her husband that are reproduced in the frontispiece of this number. The photograph was taken by White Studio, New York, during the spring of 1938 shortly before Professor Bentley's retirement, when he was 68 years old. The signature, taken from a bank check, was written in 1954, a few months before his last illness.

K. M. D.

Egon Brunswik: 1903-1955

Egon Brunswik was born in Budapest, March 18, 1903. His father was Hungarian—an engineer in the Austro-Hungarian government. His mother was Austrian. His childhood tongues were Hungarian and German. When only eight, he was sent to Vienna to be educated at the famous Gymnasium of the Theresianische Akademie as a preparation for entering government service in the Austro-Hungarian Empire; and from this early age he was practically always on his own. The boys at the Gymnasium came from all parts of the Empire and received training in science, the classics, mathematics and history. Some of the teaching was in the boys' own respective, provincial languages (Hungarian, in Brunswik's case). This meant, for example, that he studied the history of the Empire both in Hungarian and in German and early noticed the discrepancies between the two accounts. Perhaps it was this early experience which gave him his initial insight into the merely probabilistic character of one's knowledge of one's environment.

After the first world war he was sent with his sister for some months to Sweden to recover from the malnutrition of the war years. He returned and graduated from the Theresianische Akademie in 1921. He then spent two years (1921-1923) at the Vienna Technische Hochschule studying to become an engineer. He passed the first state examination required at the end of that period but then decided to transfer to the University of Vienna to study psychology. Here he worked under Karl Bühler and came under the influence of Moritz Schlick and the Vienna Circle of logical positivists. In 1926 he passed the state examination for gymnasium teachers in mathematics and physics.

After he received his Ph.D. in 1927 Brunswik became an Assistant in Bühler's *Psychologisches Institut*. He also taught concurrently for a year in a Real Gymnasium and later for several years in the Vienna Pedagogical Institute and the Vienna Volkshochschule. In 1931-32 he went for a year to be a Visiting Lecturer in the School of Education at Ankara, Turkey, where he established the first psychological laboratory. He became Privat-dozent at the University of Vienna in 1934.

In 1933-34 the present writer, during some months in Vienna, became acquainted with Brunswik. In 1935-36 Brunswik received a Rockefeller Fellowship and spent the year as Visiting Lecturer and Research Fellow in psychology at the University of California, Berkeley. In the fall of 1937,

he returned to Berkeley as Assistant Professor, where through the normal course of events he became Professor in 1947.

In 1938 he married Else Frenkel, who had been a fellow student in Vienna and also an assistant in Bühler's Institut. They were married in New York upon her arrival from Austria. They became American citizens in 1943. In 1942-43, during a sabbatical leave, Brunswik underwent a sympathectomy for hypertension. This reduced his blood pressure which had become dangerously high; but latterly, the hypertension had flared up again. In recent years, he had had greatly to restrict his social contacts, both professional and otherwise. His death on July 7, 1955, came, nonetheless, as a complete surprise and terrifying shock and sorrow to his friends, colleagues, and students.

In Brunswik's personality there was a mixture of outgoing warmth and friendliness, on the one hand, and of basic reserve, on the other. He was a delightful and courteous host. Although sometimes seemingly withdrawn in company, he was intensely interested in persons and very acute in his estimation of them. He was most generous to students and supportive of them and had a tremendous influence on them—even upon those who did only minor work under him. His was an unique, stimulating, and dedicated mind, and this was felt by all who come in contact with him. Although he had the highest intellectual and scholarly standards for himself, he had considerable understanding of and tolerance for those less gifted.

His own interests in psychology lay primarily in the fields of perception, cognition, methodology, and theory, but he was always intensely informed of, interested in, and sympathetic to the wider psychoanalytical and sociological studies of his wife, Else Frenkel-Brunswik, and he took enormous pride in her achievements.

His psychological and scientific interests can perhaps be summarized under seven somewhat overlapping and interlocking rubrics: (1) his early envisagement of perception as in the nature of a focused and more or less successful "intentionalistic attainment" (Erreichung) of environmental entities; (2) his translation of this notion of intentionalistic attainment into more general and more behavioristic terms, so that it came to cover not merely the cue-object relations in perception but also the means-end relations in instrumental behavior and his emphasis on the merely probabilistic character of both of these types of relations; (3) the growth in his thinking of the complementary doctrines of representative design and ecological validity; (4) the basic development of the two doctrines of probabilism and functionalism with an accompanying plea against too pre-

mature a concern with problems of mediation: *(5)* his analyses of the differences between perception and reasoning; *(6)* his suggestions for future investigations; and *(7)* his encyclopedic interest in the historical development of psychology.

(1) Perception as an intentionalistic attainment. In an early book, Brunswik summarized some fifteen experiments carried out by himself and his students in Vienna.[1] These were concerned with the general topic of object constancy (*Dingkonstanz*) as it was then being much studied in Europe; but Brunswik approached the topic from a somewhat new angle. He began by drawing a distinction between a quality or quantity as immediately given in experience (a *Gegebenheit*) and the same same quality or quantity as an independently measurable environmental object—a *Gegenstand.* Influenced by Brentano, he conceived of perception as a process in which the immediately given (the *Gegebenheiten*) *intended* the independently measurable objects which latter might, in perception, be more or less successfully attained (*erreicht*). Later, of course, as he became more of an objectivist, the *Gegebenheiten* or the immediately given became merely the physico-physiological processes or cues on the subjects' sensory surfaces.

Further, Brunswik pointed out that in the usual object-constancy experiment the organism can, on the one hand, be set by the instructions, or by innate propensity, to 'intend' (be focused upon) the relatively independent, *distal* object such, for example, as: inherent object-size, independent of distance; inherent object-shape, independent of angle of regard; or inherent object-brightness (albedo) independent of the momentary amount of light then falling upon the object's surface. Or, on the other hand, the subject can be set for, intend, a more artificial, *proximal* object, more closely related to some immediate partial pattern of stimulation on the organism's receptor surface. Such proximal objects would be the size of an object as projected on a plane, a given standard distance in front of the eyes; or the shape of an oblique surface as projected on a screen, perpendicular to the line of regard; or the actual intensity of light reflected from an object's surface, abstracted from any further cues indicating the amount of light being shone onto that surface from an independent light source.

Three findings appeared. (a) If left without instructions or artificial set, the independent distal object is the one which the organism seems primarily to 'intend' (to be focused upon) and to come nearest to attaining. (b) Even when instructed by the experimenter, or by itself, to achieve this independent distal object or quality, the organism's perceptual response is pulled slightly in the direction of the more immediate proximal object. (c) When set to achieve this more proximal object, however, the organism's perceptual response will be pulled (and to a greater degree) in the direction of the independent, distal object. In a word, perception in all cases tends to achieve merely some *compromise* between these two different poles of intention.

It should be pointed out here that the notion of 'intending' or 'focusing upon,' as thus treated by Brunswik, loses its subjective connotation and becomes merely a statement of the degree to which the organism's perceptual responses do tend to

[1] Egon Brunswik, *Wahrnehmung und Gegenstandswelt: Grundlegung einer Psychologie vom Gegenstand her,* 1934, 1-244.

approximate one or another purely objectively measurable environmental entity, or to compromise between them or to do both.

Finally, Brunswik expanded this concept of perceptual poles, between which the organism's achievement could fall, to new problems outside the narrow range of thing-constancy. Thus, in one experiment, the subject had to compare cards with groups of dots. The groups were made to vary both in numerosity of dots and in sizes of the individual dots, and hence in the total area covered by them (the total amount of paint required). It was found that, if the subject was instructed to equate cards on the basis of numerosity, he was pulled somewhat towards equating them on the basis of the total areas covered; whereas, if he were instructed to equate the total areas covered, he was pulled somewhat towards the pole of numerosity.

Brunswik also investigated cases of three poles—e.g. numbers, areas, and monetary values. In one experiment this was done with groups of coins. The subject could be told to equate the groups on the basis of numbers of coins in them, or on the basis of total area covered by the coins or on the basis of their total monetary value. It appeared that, whenever the subject attempted to equate on the basis of one of these three poles, his responses were deflected (compromised) in some degree by each of the other two. Thus Brunswik anticipated the more recent American studies of motivational factors in perception. But his own interest seems to have been more in measurements of the degrees of veridical perception *per se* than in the resulting motivational distortions as such.

(2) *Cue-object and means-end relations: Probabilism.* After coming into contact with American behaviorism Brunswik translated the concept of the relation of givens (*Gegebenheiten*) to objects (*Gegenstände*) in perception into the more objective terms of the relation of perceptual cues to the to-be-perceived objects (whether proximal or distal). He developed the additional notion of a similar relation of means-objects to goals on the instrumental side.[2] The organism on its perceptual side receives and selects among cues in order perceptually to attain objects, and on its instrumental side it selects and manipulates means-objects in order to reach (attain) goals. So he came to put great emphasis upon the fact that both cue-object relations and means-goal relations, to which the organism has thus to adjust, have usually only varying degrees of 'probable' applicability or validity. For example, one single cue for perceiving the albedo of a surface is simply the amount of light reflected from that surface, but under special conditions, where other cues are impoverished, the subject may be misled by this one cue and over-evaluate it. Thus, for example, it was found (Gelb) that a black disk in a dim room reflecting a large amount of light may (if the source be hidden) be seen as relatively white. Without adequate other cues the organism over-evaluates this one cue of reflected light. If however, a bit of white paper be added and the light also reflects from this really white paper, then the black disk is immediately seen as black.

Similarly, Brunswik became impressed with the fact that in the normal environment given means-objects are apt to lead to given goals with only certain degrees

---

[2] E. C. Tolman and Egon Brunswik, The organism and the casual texture of the environment, *Psychol. Rev.*, 42, 1935, 43-77.

of probability. He conceived a new sort of experiment in animal learning to illustrate this point.[3] Up to that time, in most of the orthodox instrumental learning experiments, the probability that a given means would lead to the reward had usually been set as 1.00 and the probability that an alternative means would lead to reward had been set at 0.00. This all-or-none relation was felt by Brunswik to be an artificial, non-representative situation. Specific means (as well as specific sensory cues) lead in real life to given goals (and given to-be-perceived objects) with varying degrees of probability. In the means-end case there may also enter in a second factor. In the real environment there is also a tendency for a given goal (e.g. food) to become exhausted after repeated gettings to it. Hence a given means (e.g. a given location) may become less good (principle of exhaustible supply) with repetition.

Brunswik's experiment illustrated the first point. In a simple T-maze the correctness of the one side and the incorrectness of the other were given different probabilities with different groups, such as: 100%–00%; 75%–25%; 67%–33%; 100%–50% and 50%–00%. He discovered that, whereas 67%–33% was below threshold, the other probability-combinations led to the choice of the more probable side with increasing frequency in the order 100%–50%; 75%–25%; 50%–00% and 100%–00% (the last combination being the best). A brief (inadequate) summary of these results would be that above a certain threshold the rats were governed by, and in some degree responded appropriately to, the increasing differences in probability values of the two alternative means-routes.

Several additional experiments, somewhat similar to this rat experiment but using human subjects, may be briefly mentioned. Brunswik and Herma early did an experiment in Vienna in which the subject had to judge which was the heavier of two weights, one presented to each hand.[4] In a semi-randomized fashion the weight on one side was in 67% of the paired lifts heavier than that on the other side. In a series of interspersed test-trials where both weights were equal—either both heavy or both light—the subjects showed a first increasing and then an unexplained somewhat decreasing tendency (when both weights were equal) to judge the one on the 67%-side as lighter (contrast). In other words, the contrast illusion could be set up on the basis of merely probable and not univocal relations between positions and weights.

Levin repeated the Brunwik-and-Herma experiment with similar results.[5] Jarvik did an experiment on the building up of the relative expectations of two alternative symbols attached with different probabilities to orally presented words.[6] His results, although using such different materials, were nonetheless also similar to those of Brunswik and Herma.

---

[3] Brunswik, Probability as a determiner of rat behavior, *J. exp. Psychol.*, 25, 1939, 175-197.

[4] Egon Brunswik and Hans Herma, Probability learning of perceptual cues in the establishment of a weight illusion, *ibid.*, 41, 1951, 281-290.

[5] M. M. Levin, Inconsistent cues in the establishment of perceptual illusions, this JOURNAL, 65, 1952, 517-532.

[6] M. E. Jarvik, Probability learning and a negative recency effect in the serial anticipation of alternative symbols, *J. exp. Psychol.*, 41, 1951, 291-297.

(3) Representative design and ecological validity. As time went on, Brunswik became more and more concerned with general principles of experimental design.[7] He pointed out that, in the classical and, as he saw it, over-controlled laboratory sort of experiment, many of the important, independent stimulus-determiners in a perception experiment or a learning experiment become artificially 'tied' (made to covary in an unnatural fashion) or artificially untied (prevented from covarying— some of them held artificially constant and only one or two allowed to vary). He thus became more and more impressed with the need for a wide sampling over a natural *ecological* array of total stimulus-situations and over a natural ecological array of total-means-end situations. He argued, in short, that we must substitute *representative design* for what he called the classical, nomothetic, or *systematic design*. If we wish to *generalize* how well an organism actually attains its perceptual or its instrumental goals (and this he felt was the true aim of a proper psychology), we must study the organism in its natural ecology.

As an early empirical step in this direction he devised a perception experiment in which the subject (a woman graduate student) was followed about by an experimenter off and on during four weeks and questioned in 19 sessions, every few minutes, to judge the most obvious linear dimension (vertical, horizontal or oblique) of whatever frontal surface she happened to be looking at at that moment.[8] She was asked to judge under various attitudes—the only two we shall consider here being that for distal size and that for proximal size. She was also asked to judge the distances of the given surfaces that she was looking at. It turned out that the relative goodness of her judgments of distal sizes exceeded that of her judgments of projected or proximal sizes. Also it turned out that her judgments of distance were very highly correlated with the actual objective distances. Two reasons why this experiment was important were that it first made use of the notion of truly representative sampling and, secondly, it led Brunswik to his first extended use of correlation statistics. The questions to be answered were: how highly over a wide sample of situations do a subject's perceptual estimates of size *correlate* with distal sizes and with the proximal sizes, and how highly do estimates of distances correlate with true distances. The correlations (using logarithms of sizes and distances and of estimations of sizes and distances) were 0.98, 0.85 and 0.99, respectively. There were many other important questions asked and answered which, however, are too detailed for presentation here.

Brunswik likewise initiated a series of experiments (these were first begun in Vienna) in which the subjects were asked to judge, relative to schematic faces, such social characteristics as: intelligence, mood, age, character, likeability, beauty and energy.[9] Again he used correlations as measures of the degrees to which the perceptual judgments of these social qualities were determined by the various

---

[7] Brunswik, Systematic and representative design of psychological experiments, *Univ. Calif. Press* (Berkeley), 1947, 1-60; The conceptual framework of psychology, *Int. Encyol. unified Sci.,* 1, 1952 (No. 10), 1-102; Representative design and probabilistic theory in a functional psychology, *Psychol. Rev.,* 62, 1955, 193-217.

[8] Brunswik, Distal focusing of perception: Size constancy in a representative sample of situations, *Psychol. Monogr.,* 56, 1944 (No. 254), 1-49.

[9] Brunswik, *Perception and the Representative Design of Psychological Experiments,* 1956, 1-154.

different cues presented by the faces, such as length of nose, height of forehead, distance between the eyes, etc.

It was these experiments in "social perception" which perhaps more than any others led him to arrive at the notion of "ecological validity." Actually, in the true ecological environment of human observers, what are the real correlations, if any, between such factors as length of nose, distance between the eyes, etc., and say, the intelligence of individuals having such faces? Probably they are very small. It thus became obvious that in a complete 'probabilistic functionalism' (see below) one would want to compare the experimentally obtained correlations between cues and perceived social characters with the true correlations actually existing in the environment—against, that is, the real ecological validity of such cues. Furthermore, by the use of partial and multiple correlations, it could also be discovered which cues were given by the perceptual apparatus more (or less) weight than their true ecological validity, and what cues, if any, tended by perception to be given ecologically appropriate weightings.

There should also be mentioned here a couple of experiments in which the aim was to measure objectively the ecological validity of visual cues as these occur in the normal environments of our culture. Brunswik and Kamiya, using photographs from Life Magazine of 'stills' from a moving picture, measured the degree to which the proximity (nearness together) of pairs of parallel lines in the picture could be taken as a cue that the two lines were boundaries of a single object.[10] It was found that, the greater the proximity of the two lines, the more probable it was that they indicated the boundaries of a single object. The correlations were small but significant. Brunswik and Kamiya believed that this objective probability-relationship means that proximity is a cue which the organism can learn to depend upon and that proximity is presumably not an innate, autochthonous cue, as orthodox Gestalt psychology would hold.

Seidner did a similar experiment on the ecological validities of the respective vertical positions of two points as cues for positions in the third dimension.[11] He also found a definite relationship. The higher points tended to belong to objects further away.

In short, it is to be emphasized again that the probability-values of objective ecological relations are the standards against which the probability of responses are to be evaluated.

(4) *Probabilistic functionalism.* Given the above sorts of findings and his own general point of view, Brunswik came more and more to conceive the nature of a proper psychology as being what he called a 'probabilistic functionalism.' It would be a functionalism (in the sense of American functionalism) because it would be a study of the organism's successes and failures, both in the perceptual attainment of objects and in the instrumental reaching of goals. It would also be a probabilism because in the organism's environment the objects to be perceived and the goals to be reached usually have only probable and practically never univocal relationships to their respective cues and means.

---

[10] Egon Brunswik and Joe Kamiya, Ecological cue-validity of 'proximity' and of other Gestalt factory, this JOURNAL, 66, 1953, 20-32.

[11] Brunswik, *Perception and the Representative Design of Psychological Experiments, op. cit.,* 1-154.

Furthermore, in this study of the organism's functionalistic and probabilistic achievements, he felt that at the present time a consideration of what he called "mediational problems,"[12] such as the detailed investigation of the specific sensory, motor, and cerebral activities involved should be postponed. In this connection he adopted Boring's slogan of "the psychology of the empty organism"[13] not as an ontological, but rather as a purely programmatic, proposal. We need first, he argued, to find out what the general successes of the organism's perceptual and goal-reaching attainments are and the general environmental conditions underlying these attainments, before we get involved in too detailed an investigation of the physiological or other mediational mechanisms involved.

Further, as a merely illustrative pedagogical device, he developed what he called the "lens model."[14] This was a pictorial way of suggesting that a unit of behavior (whether perceptual or instrumental) could be conceived as if it operated by a double convex lens which brought a wide-spreading array of influences issuing from one focus, through the interposition of either the perceptual or the means-end apparatus of the organism, together again at a second focus. All the rays radiating out from a to-be-perceived object in the form of an array of sensory cues are brought together to a second focus in a perceptual response. All the rays (radiating out from a goal-oriented focus within the organism in the form of possible alternative means-routes) are brought together again to a second focus upon the actual goal to-be-reached.

Finally, most importantly in his thinking was the notion of *vicarious functioning.* That is, he always emphasized the frequent equivalence and inter-substitutibility of cues and of means-routes. Different rays can be substituted one for another. But he also emphasized the fact that these cues and means form *hierarchies* in that they come to be rated by the organism in orders of relatively better and relatively poorer.

(5) Perception vs. Reasoning. Throughout both his earlier and his later studies Brunswik was intrigued by the rapid but usually only approximate attainments of perception. He envisaged perception as a primitive and relatively autonomous function within the total cognitive system of the human being. He contrasted it with reasoning, in which the subject tends either to hit the solution on the nose, or else to be all off. The differences in the distributions of errors would thus be a distinguishing criterion of whether the case was one of perception or one of reasoning. This was a concept which early fascinated him and an area of research to which he was beginning to pay much attention just prior to his death.

Thus in a recent experiment two partially overlapping groups of subjects were presented first in the laboratory with a perceptual problem in size-constancy (distal size).[15] Subsequently they were presented with the necessary data in the same situation and asked to find arithmetically the distal size of the same object at the same distance. The perceptual laboratory judgments yielded a normal

---

[12] Brunswik, The conceptual framework of psychology, *op, cit.,* 1-102.
[13] Brunswik, *loc. cit.*
[14] Brunswik, *loc. cit.*
[15] Brunswik, *Perception and the Representative Design of Psychological Experiments, op. cit.,* 1-154.

distribution about the correct distal size, with the mode close to this distal value though slightly deflected toward the proximal size. The solutions in the reasoning (paper and pencil test) yielded, on the other hand, a bimodal distribution. They tended to be either completely correct and fall exactly at the distal size or completely incorrect and fall exactly at the proximal size.

(6) Brunswik's suggestions for future investigations. As has been indicated, after he came into contact with American psychology, Brunswik conceived that the relations of means-routes to goals have, in a normal environment, an analogous probabilistic character to that of perceptual cues to objects. His own experiment with the rats, his study with Herma and the studies of Levin and Jarvik were steps in this direction. In a sense, of course, all the various studies in other laboratories under the heading of partial reinforcement bear upon the issue. But they do not usually, as Brunswik wanted, confront two or more differentially probable means-routes, one against another. What he really wanted were studies in which the alternative means-routes and their respective probabilities would correspond to their natural values in the organism's normal ecology. This neither he nor his students had yet really achieved, although several beginnings had been made. It is, of course, the difficulty of deciding what is a normal instrumental ecology for a given organism which to date has proved the major problem in the way of attempts at such experiments. Yet he was keenly desirous that such further experiments be done. For underlying Brunswik's whole thesis was a well-founded scepticism as to the *generalizability* of the laws of learning and of memory (as well as of those of perception) as we now have them as a result of our traditional, artificial, nomothetically over-controlled, laboratory conditions.

A second suggestion as to possible further work lies in the area of clinical psychology. Here Hammond and his students have made a beginning.[16] These workers have pointed out that an adequate ecological sampling means that not only subjects, but also a wide range of interviewers or projective-test-interpreters or both, must be sampled, if we are to make any true generalizations as to the ecological validity of interview results or projective-test results. They considered, for example, how well Rorschach protocols predict IQ (as defined by some independent measure such as the Wechsler-Bellevue test) and found that this could be answered only if Rorschach interpreters as well as Rorschach protocols were widely sampled. When this was done, partial and multiple correlations indicated which Rorschach factors were actually the more ecologically valid and also which Rorschach factors the various interpreters actually depended upon—which ones they weight properly, which ones they overweight and which ones they underweight. The results were illuminating both theoretically and practically.

The above are, however, only two inadequate suggestions of the types of further work which should stem from Brunswik's genius.

(7) History of psychology. Finally, Brunswik had an encyclopedic knowledge of and interest in the history of psychology and of its relations to the other sciences. He emphasized especially that psychology, while obeying all true scientific canons,

---

[16] K. R. Hammond, Environmental determinants in psychological theories, *Psychol. Bull.*, 51, 1954, 150-159; Probabilistic functioning and the clinical method, *Psychol. Rev.*, 62, 1955, 255-262.

must not slavishly follow the nomothetic thema of classical physics. In other words, while championing the unity of science in relation to its basic objectives, he also advocated diversification of content and of method. Much of his teaching (and he greatly enjoyed giving a course in the history of psychology) was devoted to the thesis that true progress in science has come from the various and imaginatively conceived different developments in the several disciplines. His interest in the 'Unity of Science' was an interest in appropriate diversities as well as in common objectives. He also believed that in the history of psychology itself he could discern specific and to some degree logically sequential trends: away from dualism, away from sensationism, away from molecularism, away from encapsulated centralism, away from nomotheticism; but towards monism, towards distal-achievementism and towards molarity—that is, towards his own doctrine of functionalism and probabilism.

This account of some of the more salient of Brunswik's concepts and their experimental implications by no means adequately portrays the complexity, the richness and the creativity of his thinking. The present report appears thin and bare when one turns from it to the perusal of Brunswik's own writings. There is a brilliance, a depth, a subtlety in his language and in his ideas which both dazzles and baffles. The reader is intrigued, stimulated, flounders, but is enormously challenged and enriched.

In the coming years, Egon Brunswik will hold an ever increasingly significant and important position in the history of psychology. His posthumous monograph[17] will help to hasten this increasing recognition. Those of us who knew and loved him can but be glad that such ever-greater recognition lies ahead, though we grieve that he did not live to see it happen.

University of California   EDWARD C. TOLMAN

---

[17] Brunswik, *Perception and The Representative Design of Psychological Experiments, op. cit.,* 1-154.

William Lowe Bryan: 1860-1955

The death of William Lowe Bryan on November 21, 1955, marks the passing of the next to the last charter member of the American Psychological Association as well as of a past president of the Association (1903).[1] The interest in psychology thus indicated was maintained by Dr. Bryan through his long life. During the many years he served as President of Indiana University, he actively fostered the subject by both personal and financial encouragement. The publication of his monograph, *On the Psychology of Learning a Life Occupation,* and his privately printed autobiographical commentary, *Adventure in Psychology, 1885-1902,* both after his retirement from the university presidency give evidence of his continuous interest in psychology.

William Lowe Bryan was born on a farm near Bloomington, Indiana, on November 11, 1860, and was educated in the local public schools. In 1884 he received an A.B. from Indiana University and two years later was awarded the A.M. His Ph.D. was earned at Clark University in 1892 with a dissertation, under G. Stanley Hall, "On the development of voluntary motor ability."[2]

Dr. Bryan's academic career was extremely versatile. Though he rendered all his academic service at Indiana University, he held appointments in various fields. In 1884 he was instructor in Greek; in 1885 he was appointed Assistant Professor of Philosophy and acting Instructor in English. In 1886 his rank was raised to Associate Professor of Philosophy; in 1887 he was made Professor of Philosophy and Instructor in Elocution. His interest in philosophy was maintained together with his later psychological interests. In collaboration with his wife, Charlotte Lowe Bryan, he published several books on Plato, with regard to problems of teaching. He also published in 1940 his lectures on *The Wars of Families of Minds.*

After obtaining his doctoral degree in psychology at Clark, Bryan re-

---

[1] Of the 26 charter members of the American Psychological Association, the only surviving member now is Lightner Witmer. The 10 who have survived longest, with their dates of death, are, in order: Lightner Witmer (living), W. L. Bryan (1955), John Dewey (1952), G. T. W. Patrick (1949), E. W. Scripture (1945), E. B. Delabarre (1945), J. McK. Cattell (1944), Joseph Jastrow (1944), W. H. Burnham (1941) and Frank Angell (1939). Scripture, whose death was not well noted at the time, died near Bristol, England, on August 5, 1945, at the age of 81 yr. For the other charter members, see Wayne Dennis and E. G. Boring, The founding of the A.P.A., *Amer. Psychologist,* 7, 1952, 95-97.—Ed.

[2] This JOURNAL, 5, 1892, 125-204.

turned to Indiana, where he established one of the first psychological laboratories in this country and headed the teaching and research. He was one of the earliest American students to go to Germany to study psychology. In the year 1886-1887 he went to Berlin and in 1900-1901 he studied in Paris and Würzburg.

The outstanding psychological research so intimately connected with Dr. Bryan's career comprises the studies he made with E. H. Lindley on arithmetical calculators, especially on Arthur F. Griffith, and the famous studies on the physiology and psychology of the telegraphic language. These studies, which Dr. Bryan regarded as subsumed under the general theme of the psychology of learning of a life occupation, interested him throughout his long career. After retiring from the presidency of Indiana University, he returned to his materials on the calculators and after rounding out his study, he published it in a monograph in 1941[3] together with a reprint of the telegraph-studies, which originally appeared before the turn of the century.[4]

To contemporary psychologists Dr. Bryan is best known for his study of the acquisition of the telegrapher's skills, performed in collaboration with Noble Harter, who was not a psychologist but a telegrapher with an interest in the scientific examination of the learning process. Of all research in the field of learning, this work is perhaps more widely known than any other single study. Its importance to psychology is best indicated by the fact that it is still reported in some detail in every important textbook which presents general psychology with a scientific orientation. Providing early illustrations of learning curves and the first demonstration of the 'plateau' (the temporary cessation of improvement in performance commonly found in learning complex tasks), this research is a psychological classic. The learning plateau and questions which Bryan raised concerning its relation to levels of complexity at which a task may be learned are still subjects of active research.

In number, Bryan's psychological publications are relatively few, but they are not comparable to the shorter journal articles of today, which are limited in scope by editorial pressure to more restricted investigations. Bryan's work provides a clear demonstration that the significance of a

---

[3] W. L. Bryan, E. H. Lindley, and Noble Harter, On the psychology of learning a life occupation, *Ind. Univ. Pub. Sci. Ser.*, 11, 1941, 1-129.
[4] Bryan and Harter, Studies in the physiology and psychology of the telegraphic language, *Psych. Rev.*, 4, 1897, 27-53; Studies on the telegraphic language: The acquisition of a hierarchy of habits, *ibid.*, 6, 1899, 345-375.

man's contribution to science is not well measured by the number of titles in his bibliography.

It was an unique merit of Dr. Bryan that his administrative career, which was aimed foremost and always at the greatest service to students, redounded effectively to the advantage of scientific research. Generations of scientists and scholars, notably his psychological colleagues, have been able to testify to his help in carrying out their work. Always he lent a sympathetic ear to the call for help in arranging schedules and the purchase of apparatus. In his own way he was able to win the gratitude and esteem of his co-workers. Thus many of them were at one with the general population of city and state in honoring him in life and mourning his loss when he died.

D. G. ELLSON

Indiana University

John Carl Flugel: 1884-1955

On August 17, 1955, 'Jack' Flugel died in London. Seventy-one years earlier, he was born, on June 13, 1884, in Liverpool, England. During these years he had travelled widely, not only geographically but also in terms of his personal and professional interests.

To appreciate Flugel's contributions to psychology it is important to realize that, during most of his professional life, he was playing two roles: one as a practicing psychoanalyst and the other as a lecturer in experimental psychology. During the six years I knew him personally, he occasionally discussed the conflicts which these two roles had produced, for he believed that such conflicts could not be resolved by accepting one role and abandoning the other. He felt strongly that the two approaches were working toward similar general goals, toward a better understanding of human relations, and he did all he could to encourage each to proceed as far and as rapidly as possible. It is a fact that these dual roles did not lead to a 'split personality,' but they undoubtedly influenced his opportunities in academic psychology.

Flugel's wide intellectual interests were encouraged during his undergraduate days at Oxford. He 'went up' at the early age of seventeen to read a 'Greats' course, became interested in experimental psychology and joined a small group of students studying with William McDougall. This interest later attracted him to Würzburg, to a laboratory strongly influenced by Külpe's approach to research on thinking. In 1909 he joined the staff at University College, London, becoming Spearman's first assistant. Spearman's general research program had just begun and Flugel was made responsible for experiments on fluctuations in attention and in mental work. He remained a member of the staff until the time of his death.

Even after he had decided upon a career in psychology and had been influenced by McDougall's experimental approach, his broad interests in human behavior persisted. They soon led him into close contact with Freudian theory and the practice of psychoanalysis so that he became one of the early champions of Freudian psychology in Britain.

Although he himself saw as his "most distinctive contribution to teaching" a course of lectures on psychoanalysis given annually over a period of almost 30 years, he continued to conduct research and to direct the work of graduate students to the end of the last academic year. In his words, he "acted as a sort of liaison officer between psychoanalysis and 'academic psychology.' " The temporal sequence of his early publications gives a clear picture of the dual role he was playing. His first book, *The*

Psychoanalytic Study of the Family (1921), reflected the important impact Freud's views had upon his early thinking. This work was followed in 1928 by a monograph, *Practice, Fatigue, and Oscillation,* that reported the results of an extensive series of carefully conducted experiments. The remainder of his main publications, including *Studies in Feeling and Desire,* which appeared posthumously in 1955, showed how his thinking had been influenced by his dual role in psychology. They emphasized particularly the conative (he used the term *orectic*) and social aspects of human behavior but did so with a definite psychoanalytic flavor. Perhaps best known generally is his *A Hundred Years of Psychology,* which, as he says in the preface, "Is deliberately constructed on more impressionistic and less systematic lines." This book, as do all his main writings, reflects the great interest he had in people and in social problems.

During this long period of service Flugel was able to watch and to participate in the major developments of British psychology. The present organization of the British Psychological Society owes much to his personal efforts. He served the Society as its Secretary, Librarian, President and, during several of the post-war years, as Chairman of its Social Psychology Section. Speaking five languages, he contributed much to the good relations between psychology in Britain and the other Western European countries.

Shortly before his death I heard a student remark: "If I could grow old just half as gracefully as Dr. Flugel, I would consider that I really had accomplished something worth while." Flugel was indeed a man of many graces. His unusual ability to handle interpersonal relations made him frequently sought as a confidant and advisor by both his students and his colleagues. As a lecturer he was superb. A final assessment of his contributions to psychology must place as much, if not more, emphasis upon his personal success in teaching and his encouragement of young psychologists as upon his research and his writing.

University College, London

Roger W. Russell

BOOK REVIEWS

The Foundation of Statistics. By L. J. SAVAGE. New York, John Wiley and Sons, 1954. Pp. xv, 294.

As asserted by its title, the book under review represents an attempt to examine, carefully and rigorously, the foundations of statistics. The examination is motivated by a *personalistic* (as opposed to an *objectivistic* and to a *necessary*) view of the interpretation of probability. Seeking a basis for this personalistic orientation, Savage considers first the general problem of decision-making in the face of uncertainty. In this endeavor he finds it important to "develop, explain, and defend a certain abstract theory of the behavior of a highly idealized person faced with uncertainty" (p. 5).

The first seven chapters are devoted to an exposition of the author's theory of behavior, and to the personalistic theory of probability and the theory of utility to which it leads. The psychologist, *qua* psychologist, will be interested to compare the properties which Savage's seven postulates give to his "highly idealized person" with the characteristics of real persons, but he will not find the going easy. Savage is primarily a mathematician and—in the opinion of the reviewer—a mathematician's mathematician. It is therefore to be expected that in spite of the skillful and polished writing which makes the 'easy' parts really delightful reading, one will not breeze through the book. The author has attempted to express his views in a form which makes a minimum of demand upon the mathematical or statistical background of the reader, but his material is certain to offer difficulty to one who has not had some considerable experience with highly abstract symbolic reasoning. It is perfectly clear that, as he wrote, Savage was keenly aware of the mathematicians looking over his shoulder—and he has not neglected them! Although "the explicit mathematical prerequisites are not great; a year of calculus would in principle be more than enough," we find allusion to Lebesque measure, Borel measure, the Hausdorff moment problem, countably additive measures *vs.* finitely additive measures, Jordon content, σ-algebras, convex sets, Fourier integral, and the like. I do not mean to imply that the author has not been successful in his endeavor "to keep the book as free from technical prerequisites as its subject matter and its restriction to a reasonable size permit," but only to hint at the hazards the book offers to the mathematically unsophisticated.

Chapters 8-17 are devoted to what Savage calls "statistics proper" and the relation of its problems and their common solutions to his theory of personalistic probability. The psychologist as a statistics-user may be interested in these chapters in which minimax theory and the theory of estimation are discussed in their relation to the personalistic interpretation of quantitative probability developed in the first seven chapters. It is my belief, however, that the psychologist will be most interested in the early chapters; and he will not be in position to profit from the later chapters without a reasonably complete understanding of the earlier ones.

The author begins by considering: (*a*) a set of "states of the world"—the

world about which decision must be made, which may be a very restricted or a very unrestricted world; (*b*) a set of consequences, apparently subjective states of the person, considered independently of the states of the world; (*c*) a set of acts, each of which maps each possible state of the world into a consequence; and (*d*) a preference relation, "is not preferred to," symbolized by "$\leq$" and obtaining between *acts*.

The first postulate is one which seems rather cavalierly to surmount a crowd of difficulties which have plagued generations of psychologists. When *real* persons attempt, under whatever motivation, to make judgments of even the simplest kind about a set of even rather simple stimulus-objects, their judgments persist in displaying an annoying lack of transitivity. Yet Postulate *1* asserts that the preference relation obtaining between *acts* for the idealized person is a simple ordering. This would seem to imply that our "person," faced with a list of acts, together with complete information concerning the nature of the mapping of each act from each of the possible states of the world into subjective consequences, is enabled to make judgments which will order the acts of the set in a preference order which may be displayed as a set of loci along a line, although no significance—or even meaning —is attached to the distances separating different local representations. This postulate of the simple ordering of the preference relation among acts make it possible for the author to define the relation "is not preferred to" as it obtains between *consequences*—subjective states of the person. It is not possible for me to determine whether the apparent artificiality of this definition arises more from my conviction that real persons make decisions about acts on the basis of their ordering (simple or otherwise) of consequences, or from the difficulty I still have in following the abstract logical argument embedded, as it must be, in a context of relatively forbidding mathematical apparatus. I believe, however, that it arises from the conviction that, for real persons, preference-ordering of consequences is a necessary prior condition to the preference-ordering of acts. At first glance it might seem that a basis for such a personalistic theory of probability as Savage seeks could be established upon a postulate of the simple-ordering property of a preference relation holding among *consequences*. Such a postulate is not, however, equal to the demand which the theory would make upon it; the simple-ordering of the preference relation among acts, or its equivalent, is required.

When the reader once thoroughly understands exactly what is asserted by Postulate *1*, he is convinced (even though he should give up the task of following the remaining steps in the derivation) that for the person who can make such a precise simple ordering of his preferences for *acts* on the basis of complete information concerning the way these acts map into subjective consequences, an extremely detailed personalistic theory of probability ought to be derivable; but this conviction is one thing, and the realization of the theory quite another. To the layman in such affairs, it seems truly remarkable that by the addition of only six postulates, most of them surprisingly simple when judged by the English translations and elaborations proved by Savage in the text, it is possible to work out the theory of quantitative personal probability in such detail as he has done.

No translation of the postulates into English can fail to do violence to them; the rigorously precise formulations of mathematical symbolism find expression in the language of everyday life only in extremely verbose forms. Yet most of us will be tempted to a more careful examination of the theory of behavior which Savage

presents only by brief (and consequently distorted) English translations of the postulates which underlie the theory. I present them, therefore, but with the qualification noted. The first already has been discussed. The second asserts that if our person does not prefer one of two acts knowing that a certain event obtains, and if he does not prefer that act knowing that it is not the case that the same event obtains, then he does not prefer the act. This Savage appropriately names the "Sure-Thing Principle." The first two postulates permit the definition of a *preference relation among consequences* in terms of the preference among acts. The third postulate then asserts that the knowledge that a particular event obtains cannot establish a new preference among consequences nor, in general, reduce a previous preference to indifference. Postulates 4 and 5 make possible a definition of *qualitative probability*—a relation obtaining among *events*—in terms of the preference relation among consequences. Postulate 4 asserts that if our person would elect to stake any prize upon one rather than upon the other of two events, he would make the same election whatever the value of the prize; while Postulate 5 asserts that there is, for our person, at least one worthwhile prize in this situation. To quantify the personal probability, Postulate 6 is required. It asserts the possibility of partitioning the set of states of the world into an arbitrarily large class of subsets—mutually exclusive and exhaustive of the set of possible states of the world—in such a way that each subset is approximately equally probable with each other one. A final, seventh postulate is necessary to extend the concept of utility—a concept intimately linked with that of quantitative personal probability—to acts with an infinite number of consequences.

The book is not easy! It will have a limited appeal to psychologists, but it may be read with profit by those who have begun to develop an interest in the general problems of decision-making, who are sympathetic to mathematical models, and who have the fortitude to stay with this difficult but capable presentation.

University of California   R. F. JARRETT

Transformation Analysis of Factorial Data and Other New Analytical Methods of Differential Psychology With Their Application to Thurstone's Basic Studies. By YRJO AHMAVAARA. Helsinki, Academia Scientiarum Fennica, 1954. Pp. 150.

In this short book Ahmavaara has at one place or another tackled virtually the entire methodological side of factor analysis. He amply illustrates his own new methods and his opinions regarding unsolved technical questions with the classical studies of the *Primary Mental Abilities* (*PMA*) by the Thurstones. He has also attempted to *interpret PMA* in the light of his methodological contributions; this book is psychological as well as mathematical.

Part I deals with Ahmavaara's "transformation method" which "purports to answer the question: *How is it possible to make exact comparisons between different factorial studies carried out with different experimental populations?"*— when the two test-batteries under consideration contain at least some tests in common. What Ahmavaara (and Thurstone) mean by "experimental population" remains a mystery; apparently this unconventional expression may be translated as either *sample* or *population* in modern statistical terminology. Chapter I of Part I develops the mathematical theory of the method which rests ultimately on Pearson's theory of the effect of natural selection. This chapter probably involves a heavier dose of concise matrix algebra and advanced calculus than most psychologists would

care for, but the entire following chapter is given to a very carefully worked-out example. Consequently, a non-mathematician may apply the method, cookbook style.

Is Ahmavaara's transformation method worth the effort of its rather heavy computations? It surely is, for it does provide *an* answer to the most important question it proposes to answer; an Ahmavaara transformation yields a "comparison matrix" whose elements "indicate the cosines of the angle between the factors of different studies." It is thus possible to assess quantitatively the extent to which 'dominance,' say, found in a study in Indiana is 'dominance' in another study from Pakistan.

Chapter III of Part II applies the transformation method to Thurstone's original adult *PMA* compared with an analysis on different adult subjects of a sub-battery of the original tests. A similar application is made to the later *Factorial Studies of Intelligence*, where the subjects are fourteen-year-olds. The latter study yields but three factors (Word-fluency, Space, and Verbal comprehension) which Ahmavaara deems invariant according to his "invariance coefficient," while *none* of the earlier *PMA* factors is considered invariant. In Chapter IV of Part II Ahmavaara attempts to explain this lack of invariance in adult ability factors with a line of reasoning which is demonstrably false. Briefly, he applies the "general law that the intercorrelations of ability tests are never negative"—which is not a law but a crude generalization—and the "old truth that the intercorrelations of ability tests decrease along with the increasing age of subjects" to deduce that, as age increases, "the limiting case of an ability configuration is a rectangular spherical cone" which implies that "any corner-shaped projections (primary factors) at its border are bound to become more difficult to ascertain." No *a priori* argument can guarantee the presence or absence of tightly clustered test-vectors, whether on the border of a rectangular spherical cone or not. Nor does the argument of small, necessarily positive correlations militate against clearly defined "corner-shaped projections"; the angular separation of test-vectors in the common factor space is as much a function of their communalities as their intercorrelations. A more plausible explanation of the lack of invariance in Thurstone's adult *PMA* factors is sampling error. The original *PMA* study was carried out on a small sample which was rendered virtually non-existent by the use of frightfully inefficient (in a statistical sense) tetrachoric coefficients of correlation.

Chapters V, VI, and VII are devoted to the development of the notion of "residual spectra." Using four rather *ad hoc* methods of evaluating residual co-variances, Ahmavaara postulates what he calls "latent common factors." In spite of the new terminology, these ideas do not seem so original as one might be led to believe. Anyone who has ever done a factor analysis is confronted with the problem of "when to stop factoring." Should one extract "too many" and strain at interpreting the subsidiary factors, or should one quit early and wonder what is going on in residual outer space?

More interesting is Ahmavaara's notion of the "differentiation process of the factors" into their "fine structure components." Again we have a common phenomenon dressed up in new language, but a good demonstration of the decomposition of supposedly primary factors is provided. It seems as if there is an historical precedent for Ahmavaara's decomposition of *PMA: PMA*'s decomposition of Spearman's *g*.

On the technical side, Part III is marred by Ahmavaara's confusion regarding "ideas expounded by Godfrey Thomson, who has contended that in factor analysis

one is arbitrarily maximizing the specifics and, consequently, minimizing the communalities and the number of common factors." He maintains that "Thomson's contention has consequently been proved erroneous as a mathematical theorem of factor analysis," citing the definitive paper of Ledermann. A more careful reading of Ledermann shows that exceptions to Thomson's contention are mathematical curiosities (as Thomson well knew) and that we invariably do commit ourselves to minimizing the test variance analyzed in the common factor space.

To summarize, Ahmavaara has in Part I presented his fundamental contribution, the transformation method. This method is probably the best extant for scientifically adjudicating the most important psychological question in factor analysis: are factors true, invariant psychological entities, or are they merely convenient descriptive variables? Chapter III applies the method in a straight-forward manner to the classic *PMA* studies. It is regrettable that the last four chapters appear with the first three, for they only dilute the impact of the earlier, very significant contributions. In a sense this book is a study in contradiction. The last half, in which Ahmavaara often engages in wild flights of speculation, contains just the sort of material which Ahmavaara's transformation method seeks to put out of business.

University of California   HENRY F. KAISER
University of Southern California   WILLIAM B. MICHAEL

Sovereign Reason and Other Studies in the Philosophy of Science. By ERNEST NAGEL. Glencoe, Illinois, The Free Press, 1954. Pp. 315.

"By making manifest the nature of scientific reason and the grounds for a continued confidence in it, contemporary philosophy of science has been a servant of men's noblest and most relevant ideals." The spirit expressed in this sentence pervades the sixteen essays in this collection which exhibits the character and temper of American philosophy at its best. Although they were published separately between 1936 and 1950, the essays are held together by a concern with the problems of philosophy of science. The collection does not claim to constitute a systematic discussion of those problems. All save one are critical studies of "contemporary philosophers who have been occupied . . . with the content or method of modern science, and who have thereby sought to illumine the nature of human reason." The basis for the evaluation of their work is contained in four major themes. The most persistent and central issue is the relation of theory to concrete experience. The development of mathematical physics has made the concern with this type of problem acute, and the character of current theories which possess aspects apparently "alien" to experience has enhanced its significance. A second major theme is the concern with the reliability of knowledge, what role sensation, insight, or belief play in achieving knowledge, how probable inferences are to be interpreted, and how the validity of evidence is to be determined. A third theme is "the perennial quest for a total view of the universe." Questions arising here are related to most of the issues of traditional metaphysics. The fourth theme deals with the relations between science and society. Here we meet the problems of social determinants and consequences of science, and the relations of scientific findings to individual and social values.

These themes, recurrent in the history of thought, are of concern not merely to philosophers but also to scientists themselves, some of whom, such as Helmholtz, Mach, Hertz, and Poincaré, have contributed the most penetrating and clarifying analyses of scientific thinking. Yet, scientists and philosophers also have frequently

obscured the fundamental issues, and the best feature of the book is the illuminating criticisms to which Nagel subjects the work of his fellow-philosophers whose positive contributions he stresses and appreciates. He shows by their examples what not to do in the philosophy of science. Russell and Whitehead are mentioned as having ignored the concrete ways in which certain scientific abstractions are to be employed. It is not fruitful to indulge in a radical scepticism concerning the possibility of ever making correct judgments and to raise the problem of the existence of an external world as Reichenbach, Russell, and others have done, although it is necessary to ask what the conditions are under which errors of judgment and conclusion are committed. A verbal game has replaced the search for genuine knowledge when a quest for a total view of nature terminates only in a set of ultimate categories which are compatible with every course of events. The basic categories constructed by Peirce, Whitehead, or Dewey still have to prove their worth as instruments of scientific thinking. Of special interest to social psychologists should be Nagel's insistence that the context in which the character of sub-atomic events is described is not the context in which human values are operating.

Nagel's central position is that of "contextual naturalism," a "this-worldly" philosophy, devoid of romantic sentimentalism, and "America's most significant contribution to philosophic intelligence." Its historic roots are in the pragmatism of Peirce, James, Mead, and Dewey, but it is not so to be identified. It will probably never be a finished edifice. Its central thesis is the essentially incomplete and fundamentally plural character of existence with qualitative discontinuities and loose conjunctions. "Contextual naturalists exhibit a profound distrust of philosophic systems which attempt to catch once for all the variegated contents of the world in a web of dialectical necessity." Contextual naturalism emphasizes the conditions for the occurrence and properties of any event whatsoever. Thus, quality is an objective constituent of nature. Contextual naturalism is vigorously anti-reductionist, maintaining that "the world contains at least as many qualitatively distinct features as are disclosed in human experience, and not a fewer number of them." It does not conceive the primary moral problem to be that of discovering a fixed set of ethical norms since moral problems are plural in number and specific in character. Contextual naturalists confidently extend the use of the methods of science to human problems, knowing that these methods provide no guarantee against errors, do not preclude alternative solutions, and do not terminate in eternally valid conclusions. Contextual naturalism "expresses the convictions of a people confident that a bold but disciplined intelligence is still a creative power in the world."

University of Colorado   KARL F. MUENZINGER

Bibliography on Hearing. Prepared by the PSYCHO-ACOUSTIC LABORATORY OF HARVARD UNIVERSITY; S. S. STEVENS, Director; J. G. C. LORING, Compiler, and DOROTHY COHEN, Technical Editor. Cambridge, Harvard University Press, 1955. Pp. 599.

This is a second edition of the *Bibliography in Audition* which was prepared by the staff of Psycho-Acoustic Laboratory in 1950. The first edition consisted of two paper-bound, mimeographed volumes listing somewhat over 5,000 titles. Most of the titles in the first edition were for the period 1939-1947, although some were of earlier date, and a few titles as late as 1949 were included. The format of the second edition—a single volume, printed and bound in hard covers—is both more attractive

and more convenient. The *Bibliography* now contains about 11,000 titles, including all of the titles of the first edition. It has been brought up to date through 1952, and earlier titles also have been added.

The main section of the *Bibliography* is the list of titles, alphabetized by author. A wide range of the world literature has been searched; a sample indicates that about one-third of the articles were written in languages other than English. There have been some changes in emphasis since the earlier edition. According to the preface, "the present volume devotes more space to the psychology and the acoustics of music, to deafness and the deafened, to ultrasonics, and to the effects of drugs on human and animal hearing. New fields of inquiry, such as information theory, have been added." The form of citation is similar to, but not identical with, that used in publications of the American Psychological Association. Foreign titles are usually given in the original, and in all cases they are translated into English.

In addition to listing the titles, the *Bibliography* provides some aid to the reader in finding the titles which may be of special interest to him. This is, of course, the chief purpose of such a compilation, but here the *Bibliography* leaves something to be desired. The reader is directed first to scan a list of 315 topics to find the classification of the subject in which he is interested. (This is a somewhat smaller number of categories than was used in the first edition.) The instructions continue: "Having found the topic, note its serial number, and then turn to where a group of names is listed under this number (pp. 561-591). These are the names of persons listed as first authors on papers that deal with the subject. The reader is left, of course, to determine which of the papers by a given author are relevant to his interests." Since many authors are represented by a score or more of titles, this task may be a time-consuming one. Often the reader is also left to determine which of several authors of the same surname is the relevant one. In the second edition, as in the first, initials are used only sporadically in the classification to distinguish different authors of the same surname. Various systems might have been used conveniently to give unambiguous references to the titles, thus increasing substantially the value of the classification by subject.

Most readers are sure to find that the *Bibliography* will supply titles new to them even in the fields of their own special interests. On the other hand, readers with special interests will probably wonder why certain titles were omitted. For example, those interested in auditory localization may note the absence of titles such as these: M. D. Arnoult, Localization of sound during rotation of the visual environment, this JOURNAL, 65, 1952, 48-58; C. T. Morgan and Elliot Stellar, *Physiological Psychology*, 1950, 1-609 (the first edition, 1943, by Morgan is cited); Henri Piéron, L'orientation auditive latérale, *Année Psychol.*, 23, 1922, 186-213; H. A. Witkin, S. Wapner, and T. Leventhal, Sound localization with conflicting visual and auditory clues, *J. Exper. Psychol.*, 43, 1952, 58-67; P. A. Zahl (Editor), *Blindness: Modern Approaches to the Unseen Environment*, 1950, 1-576. Readers might well inform the Psycho-Acoustic Laboratory of other omissions that they discover, to the end that a succeeding edition of the *Bibliography* may be even more nearly complete.

Students of audition are greatly indebted to the staff of the Psycho-Acoustic Laboratory for this painstaking work; it will be a major aid in utilizing the already extensive and rapidly expanding literature of the field.

University of California   MARK R. ROSENZWEIG

Survey of Clinical Practice in Psychology. Edited by Eli A. Rubinstein and Maurice Lorr. New York, International Universities Press, 1954. Pp. xvii, 363.

Calling upon 32 contributors specializing in one or another area of clinical psychology, the editors have compiled 27 authoritative chapters on the work of clinical psychologists in as many different locales. Part I covers long-established centers of activity, such as university clinics, institutes for juvenile research, institutions for the mentally defective, state hospitals, and community clinics. Part II describes large-scale government programs, such as the Veterans Administration, the Public Health Service, and the Armed Forces. Part III deals with settings which only more recently have come to include the clinical psychologist. With occasional exceptions, each chapter provides a brief historical introduction and ends with a section devoted to the future. Most chapters include sections on diagnosis, psychotherapy, and research. Unfortunately, however, the average chapter is only about 12 pages in length, and little opportunity is provided for thorough discussion of any one topic.

The book is an impressive testimonial to the widespread acceptance of the clinical psychologist in the United States. A most striking contrast would be provided by an attempt to locate similar psychologists in any European country at the present time. Occasional counterparts could undoubtedly be found, but there is little doubt that clinical psychology is a distinctly American development. From the first decade of the present century, diagnostic testing has been the principal work of the clinical psychologist. This book leaves little doubt that psychotherapy is now also a routinely accepted activity. Even reading between the lines fails to reveal any great concern by the psychologists themselves over the propriety of their role in therapy. That role seems limited more by lack of time than by opposition from other disciplines. Although each chapter describes ongoing research, there is little conviction that research constitutes any major portion of the work performed by clinical psychologists in most settings. There are, of course, some exceptions. Heiser's chapter on the Vineland Training School for mental defectives is a concise description of a long-established research 'laboratory' which also has service functions. The account by Waldman and Phillips of Worcester State Hospital highlights the research carried on there. Hildreth's chapter on the massive program of the Veterans Administration illustrates the care which has been taken by that enormous federal agency to promote research along with service.

Despite the overwhelming assurance contained in almost every chapter that clinical psychology is here to stay, the portions of each chapter devoted to the "Future" present a generally sober outlook. Few writers seem to be bursting with optimism despite the fact that each one has undoubtedly witnessed personally the amazing growth of clinical psychology in the last decade. Perhaps only an overview of the entire field provides the emerging perspective of clinical psychology as a postwar profession which not even the most optimistic prewar psychologist could have envisioned. Although clinical psychologists have recently become interested in public health, or community programs as distinct from community clinics, there is little evidence in this book that psychologists are prepared to emerge from the clinics and hospitals. Two possible exceptions can be found in Brewer's discussion of the Wichita Guidance Center and Zuckerman's discussion of statewide Youth Authorities.

Since most of the chapters purposefully focus on a description of current activities in case-study fashion, the book itself is likely to become dated rather quickly. None-

theless, the editors have provided a most readable reference-work for anyone wanting a quick survey of the way in which clinical psychology has penetrated into the social fabric of modern America. It is likely that most clinical psychologists themselves would be somewhat surprised by the range and extent of their colleagues' activities.

University of Colorado   VICTOR RAIMY

Education of Mentally Handicapped Children. By J. E. WALLACE WALLIN. New York, Harper and Brothers, 1955. Pp. xiii, 485.

Education is experiencing a rebirth of interest in all types of exceptional children. The upsurge of interest seen between 1910 and 1920 with the advent of psychometric awareness is being repeated because of 'the parent movement.' Public-school enthusiasm has shifted from discovery and pilot efforts to cover-up and patronizing makeshift. A few large urban centers have provided excellent services, but schoolmen generally can no longer evade the issues. Education for *all* children is becoming a realistic necessity regardless of mental, physical, or social barriers to instruction and learning.

Out of the versatile richness of a life-time of study and experience which covers the whole era of the modern study and management of mentally handicapped children, Wallin distills a comprehensive textbook. The author describes it as "restricted to a detailed consideration of the multiple problems affecting the training and education, in public-school special classes and regular grades and in residential institutions, of the simple or nonclinical group of mentally handicapped children. It also supplies a brief overview of the socio-economic problems of vital import to the state that they present" (p. xi). The editor (H. H. Remmers) more explicitly states that herein "the reader is given a brief, concise historical orientation; an exposition of individual needs as the frame of reference for organizing education both in special and regular classes; the guide lines for effective administration; general and specific methods of instruction; the necessary characteristics and qualifications of the effective teacher; the objectives of education for the kind of children in question; the curriculum designed to achieve the objectives; statesmanlike discussion of mental deficiency as a social problem and the prevention and amelioration of mental deficiency; and last but by no means least for the serious student, copious and relevant bibliographies for further reading" (p. ix).

This is a large order and it is ably filled, but the very extent of the coverage and perhaps the established standing of the author induce a certain sweepingness of exposition which the more informed reader may find somewhat too dogmatic. Without denying the merits of the historically-sound content, younger 'authorities' will sense nostalgic overtones not in keeping with recent standpoints in a now rapidly changing field. This Odyssey-like style and treatment frustrates the very admiration it engenders. It suggests the ending of an era without anticipating the transitional exigencies of the new day. Perhaps one should not require yesterday, today, and tomorrow in a single work of such broad scope, but in a burgeoning stage *becoming* ranks with *being* and *was*. In this view this volume seems rather more 'dated' than ongoing. If such appraisal seems a bit *ex cathedra* so also does much of the book's exposition in which the very authority of statement often seems to represent only personal judgment or experience.

Bellingham, Washington   EDGAR A. DOLL

Untersuchungen zur Onomatopoiie. 1. Teil: Die sprachpsychologischen Versuche.
By HEINZ WISSEMANN. Heidelberg, Carl Winter, 1954. Pp. 241.

In this monograph the author makes an important and original contribution to
the field of language—an experimental investigation of onomatopoeia. From a
hidden source, a number of different noises were presented to Ss whose task was to
"name" each sound pattern. In one group of experiments, S was to choose a name
from several alternative nonsense-combinations of letters presented in writing, and
to improve, if possible, on one or more of these alternatives until a satisfactory name
resulted. In another set of experiments, S coined words naming the noises with no
suggestions having been provided by E. The Ss, all university educated and with
considerable verbal facility, were encouraged to make observations and to comment
on their own formulations. The protocols containing the progression from the first
attempts to those names which the Ss finally found satisfactory are reproduced in
the book.

The results of these experiments are subjected to a very careful analysis directed
towards the identification of the sources of the newly invented names. Although
closely and reciprocally interrelated, two problems can be distinguished at the out-
set: the perception of the sound pattern, and its reconstitution in speech, *i.e.* its
"name." The former depends not only on the physical stimulus but also on the per-
ceptual field. The perceptual field has both language and non-language components
which in turn play important roles in the reconstitution of the sound patterns in
speech. Given the experimental conditions, the perceived noise is elaborated in terms
of its imagined source, its structure may be translated into a visual pattern, and so
forth. Thereupon its speech-rendition is attempted, either specifically as a sound
imitation—as the traditional theories would have it—or, more typically, as a re-
production of certain structural characteristics of the pattern. The author considers
in some detail specific linguistic changes demonstrated in successive naming attempts,
and the reciprocal influence of naming and perception under conditions of repeti-
tion.

This painstaking analysis merits close study by linguists and psychologists in-
terested in language, especially since it deals with a problem which other recent
experimenters, *e.g.* Werner, have tended to leave aside. It is not a book for the
casually interested reader who is attracted by the word "onomatopoeia" in the title
and reminded of colorful expressions in poetry. There is no place for flair in this
thorough presentation, which is often difficult and involved in formulation—even
by standards of the German language.   MARIANNE L. SIMMEL

    College of Medicine,
    University of Illinois

Interviewing in Social Research. By HERBERT H. HYMAN, with WILLIAM J.
COBB, JACOB J. FELDMAN, CLYDE W. HART, and CHARLES HERBERT STEMBER.
Chicago, University of Chicago Press, 1954. Pp. xvi, 415.

In 1947, the national Opinion Research Center (NORC) was commissioned by
the Joint Committee of the Social Science Research Council and the National
Research Council on the Measurement of Opinions, Attitudes, and Consumer Wants
to study (1) "sources of error in research that depends upon interviewing as a
method of data collection" and (2) methods for controlling and reducing inter-
view-error. A series of studies concerned with sources of error in the survey-

interview, lasting nearly six years, was carried out by *NORC* with the financial support of the Rockefeller Foundation. This book reports the results of that extensive program of research in the context of the body of related work reported by other investigators. One cannot quarrel with the decision of the authors to restrict their research upon interview-error to the survey-interview; to study the interview as a research method in the social sciences would clearly be an impossible task for any one group. One may, however, question the suitability of the title chosen for their report. The applicability of findings based upon the survey-interview to other interview-settings in the social sciences and to other types of interviewer may be limited.

As a study of the survey-interview, however, this is a major work, which challenges and overthrows a number of traditional conceptions. The authors took a fresh look at the survey-interview in terms of phenomenological accounts of the interview-situation from interviewers and respondents, and the hypotheses which this approach suggested led to a number of studies which have yielded conclusions of first importance for the improvement of survey-interviewing. (*1*) The authors examined in a series of investigations the traditional and widely held belief that the communication of the interviewer's attitudes to the respondent is the major source of interviewer-error. Ideological bias was not found to effect survey-results. This is a negative finding of great importance. (*2*) Perceptual and cognitive processes in the interviewer were discovered to be major sources of error in the interview. In a number of important studies, the authors have demonstrated that attitude-structive expectations (the assumption that attitudes are unified, organized structures) and role-expectations (the assumption that respondents of given group-memberships hold certain attitudes) have significant biasing effects. These processes have hitherto been overlooked as sources of interviewer-error. (*3*) Free-answer or open-end questions were found, in several studies, to be subject to considerable interviewer-error, arising from differences in probing behavior among interviewers and from selective recording of responses. This finding indicates the need for careful control of the interviewing process when open-end questions are used.

The authors note that it will "take time and research to develop the implications of their study for error control." Certainly a first effect of this work should be to focus in the training of interviewers upon those cognitive processes in the interviewer which the authors have found to be of major importance as sources of error.

University of California   E. L. BALLACHEY

The Process and Effects of Mass Communications. Edited by WILBUR SCHRAMM. Urbana, University of Illinois Press, 1954. Pp. 586.

This book of readings originated in a need for source-materials to be used in the training of employees of the United States Information Agency and was published in the belief that it might be of wider interest. The various sections, each containing from one to nine articles, deal with the following topics: the process of communication, the primary effect—attention, the audiences of mass communication, the effect of different channels, getting the meaning understood, modifying attitude and opinion, group-effects, and special problems in international communication. There are, of course, a number of books such as this, and their contents overlap to a considerable extent. Authors like Berelson, Lazarsfeld, Doob, Cantril,

Hovland, Merton, and Lasswell, well represented elsewhere, also are reprinted here, with the same selections often being used. The book does, nevertheless, provide a representative collection of readings for the intermediate student, and the selections are reproduced at length great enough to give the student something to chew on. From the point of view of the psychologist, more experimental data might have been provided. (Two experimental papers by Hovland and his collaborators are reprinted; a modified version of a recent paper by Osgood and Tannenbaum on the principle of congruity in attitude-change is included, but the data presented in the original article are omitted.) On the whole, the book will be of greater interest to sociologists and political scientists than to psychologists.

Cornell University

RICHARD D. WALK

Behavior Theory and Social Science. By FRANK A. LOGAN, DAVID L. OLMSTED, BURTON S. ROSNER, RICHARD D. SCHWARTZ, and CARL M. STEVENS. New Haven, Yale University Press, 1955. Pp. x, 188.

This volume is described as "a terminal-year project of the authors' three years of close association while Ford fellows in behavior science at the Institute for Human Relations of Yale University" (p. ix). It deals—at about the level of the graduate seminar in social science—with a variety of topics such as theory-construction, the Hullian system (which supplies the principal orientation), definition of response, motivation, reinforcement, language, problem-solving, "free-behavior," social interaction, and culture. The departmental duplicating machine ordinarily provides for such essays all the publication they can stand.

Dublin

T. S. KILLARNEY

Stuttering in Children and Adults: Thirty Years of Research at the University of Iowa. Edited by WENDELL JOHNSON with the assistance of RALPH R. LEUTENEG-GER. Minneapolis, University of Minnesota Press, 1955. Pp. xviii, 472.

Whenever the editor of the present volume has something to say about stuttering it is important, and the book here reviewed is no exception. It is devoted to thirty years of research at the University of Iowa on the symptom-complex called stuttering. Probably no one has devoted a greater part of an active and prolific life in the professional study of a single subject than Johnson has to the subject of stuttering. His thinking, his guidance for other researchers, and his own constant search for scientific information shine through every page. The book conveys a sense of coming into the middle of something—where much has been done but nothing completed, where the conclusions reached are regarded as "relatively rough theory" rather than as finished product—which points up the scientific attitude dominating the work.

The list of contributors to the book and the list of credits to researchers is lengthly and sounds much like a roll-call of the significant students of stuttering of the past three decades. Of special interest to psychologists will be the research dealing with parental attitudes and adjustment, the variability of stuttering, the personal adjustments of stutterers, and the systematic approach to the psychology of stuttering. It is noteworthy that no attempt at summation of the research reported is attempted—if any conclusions are to be drawn they appear in Johnson's own tentative theory of stuttering. He states quite succinctly that "Stuttering appears to be an anxiety-motivated avoidant response that becomes 'conditioned' to the cues or stimuli associated with its occurrences." This is the position of the leading

authority on stuttering in the country after his own extensive research plus the values obtained by and from the research here reported.

The book is well written, carefully edited, and logically organized. It should be of considerable interest to psychologists. This reviewer regrets only that it was necessary for the editors to ignore such factors as "emotional conditions and social morbidity" for fear of being carried too far afield. Their aim was to "raise certain crucial questions which must be answered by anyone who would attempt an interpretation of stuttering from any viewpoint." This goal was successfully achieved.

University of Chicago   JOSEPH M. WEPMAN

Personnel and Industrial Psychology. By EDWIN E. GHISELLI and CLARENCE W. BROWN. Second edition. New York, McGraw-Hill Book Company, 1955. Pp. ix, 492.

This is a book I would adopt if I were teaching a course in industrial psychology at the sophomore level. Consisting of 15 well-documented chapters, it provides one of the best summaries of the field this reviewer has seen since the end of the World War II. The authors write with a sense of having grown up in the field, having seen the fads pass while the solid, established facts mature and ripen. Their perspective is well illustrated in a discussion of the measurement of job proficiency, in which all the usual methods and considerations are well described, and which concludes with a short but realistic description of the manner in which the textbook-methods are compromised with social values when managers come to grips with the question of "What is a fair day's work?"

This second edition carries a greater emphasis upon social and motivational factors than did the first. These are the areas most recently studied by industrial psychologists and in which they are likely to make increasingly important contributions as time goes on. A portion of one chapter is devoted to aspects of equipment design most of which are drawn from military research conducted during World War II; through no fault of the authors, human engineering studies in a civilian setting are notable by their absence.

Like one or two other recent books in the field, this one allots considerable space to the question of personnel classification. In a booming economy where manpower is in short supply, it is understandable that psychologists should turn attention to placement while formerly we were primarily preoccupied with selection. Unfortunately, however, the discussion of personnel classification in this book raises more questions than it answers. The authors' statement, for instance, that the aim of classification is placement of individuals in jobs to achieve "maximum effectiveness of the total organization" may provoke considerable discussion.

The chapters on interviewing, personal data analysis, and testing are written in sufficient detail to raise some of the knotty problems of research. They fail, however, to carry the reader into the practical difficulties of some fairly common operating situations—those, for example, in which criteria of performance are less than ideal or missing entirely. Admittedly, a college course ought to present 'textbook' solutions to problems of selection, but it also should steel the novice to the fact that in the bustle of business a testing program may be of some usefulness even when poor criteria are the best available.

A somewhat lengthy discussion is presented of the fatigue concept and its limitations. The notion of effective work is explored from a number of different

angles and goes beyond the narrow confines of input-output concepts to include a commendable discussion of psychological and social factors. This chapter as well as its successor on conditions of work are limited almost entirely to studies of routine repetitive tasks. As a class these are the tasks which will evaporate with the advance of automation. I had the feeling as I read these sections that they were of decreasing relevance to American industry. The industrial psychologist of the next decade had better be prepared with some notions about the effects of illumination, noise, ventilation and like on machine-tending, decision-making, and trouble-shooting rather than on the assembly of bolts, washers, and nuts.

Industrial psychologists have a curious fascination for time-and-motion study in spite of the fact that industrial engineers have been largely responsible for its development and application. Ghiselli and Brown make the usual bow in the direction of these efforts but go on to present one of the best balanced criticisms of them that has appeared in any comparable text.

Training in industry is a topic about which psychologists should have much to say that is realistic and practical. For a half century we have made a major effort to understand the basic conditions of learning. This book does a workmanlike job in presenting those facts and principles which should be applicable to industrial training, yet the discussion seems a bit ivy-covered, lacking the flavor of managerial problems. Our contributions in this area are at the technical rather than the professional or policy level. The familiar chestnuts of spaced vs. massed practice, transfer effects, overlearning, and forgetting are all well handled, but these topics fail to excite management people, I suspect, because they are not shown to be relevant to any decisions which management must make about training programs. Business and industrial managers are much concerned today about training problems, though largely at the supervisory and managerial levels. On these questions we have little to offer in published researches of a tight, clean, well-reasoned nature. The deficiencies of the text at hand are the deficiencies of research.

From this reviewer's standpoint one of the awkward features of this commendable book is the absence of chapter summaries, a lack that is especially noticeable in the final two chapters on motivation and social factors, in which there is a tendency to report serially a number of related studies without drawing them together. Perhaps the authors intended that the students should draw their own conclusions—or, more likely, that the instructor should do the summarizing.

The final proof of a textbook can only come through teaching with it. Those who use this one should not be disappointed.

Institute for Research in Human Relations   F. K. BERRIEN

Psychology for Law Enforcement Officers. By GEORGE J. DUDYCHA and COLLABORATORS. Springfield, Charles C Thomas, 1955. Pp. xii, 404.

This volume, directed as it is to a special interest, raises the question of the propriety of writing psychologies for nurses, executives, mothers, and other narrowly defined groups. If psychology is psychology, there should be no need of a separate book for each vocational minority.

The present work, which apparently is intended as an introductory text for policemen, sheriffs, detectives, and the like, may be criticized on a number of counts. First, the different chapters vary widely with respect to the amount of preparation required. Some, like "What is Psychology?" and "The Psychology of Human

Relations," are clearly directed to the beginner, while others, like "Mental Abnormality and Crime" or "Psychology of the Adult Criminal," presupposes considerable acquaintance with personality theory. (The latter chapter, interestingly enough, follows the Freudian metaphor throughout and leaves the reader with the impression that the criminal can be understood only in psychoanalytic terms. The three-line footnote by the editor which indicates that there are other points of view does not, in this reviewer's opinion, absolve him of the responsibility of providing a more representative treatment.) Second, the book seems to lack the integration that a compendium of this kind needs. It is a mixture of apparently unrelated subject-matters. The great variation among the chapters in scope and in style—from discursive accounts to taut, formal essays—is disconcerting. Third, while the quality of the contents is in general adequate, some of the writers seem merely to be summarizing what they themselves have read, lacking primary experience.

This book is not to be recommended for the purpose of introducing law enforcement officers to psychology. It would be much better to use a standard textbook at an appropriate level of difficulty along with a textbook in criminology. Such a procedure would bring to the student a much broader understanding both of psychology and of criminology.

University of Chicago   RAYMOND J. CORSINI
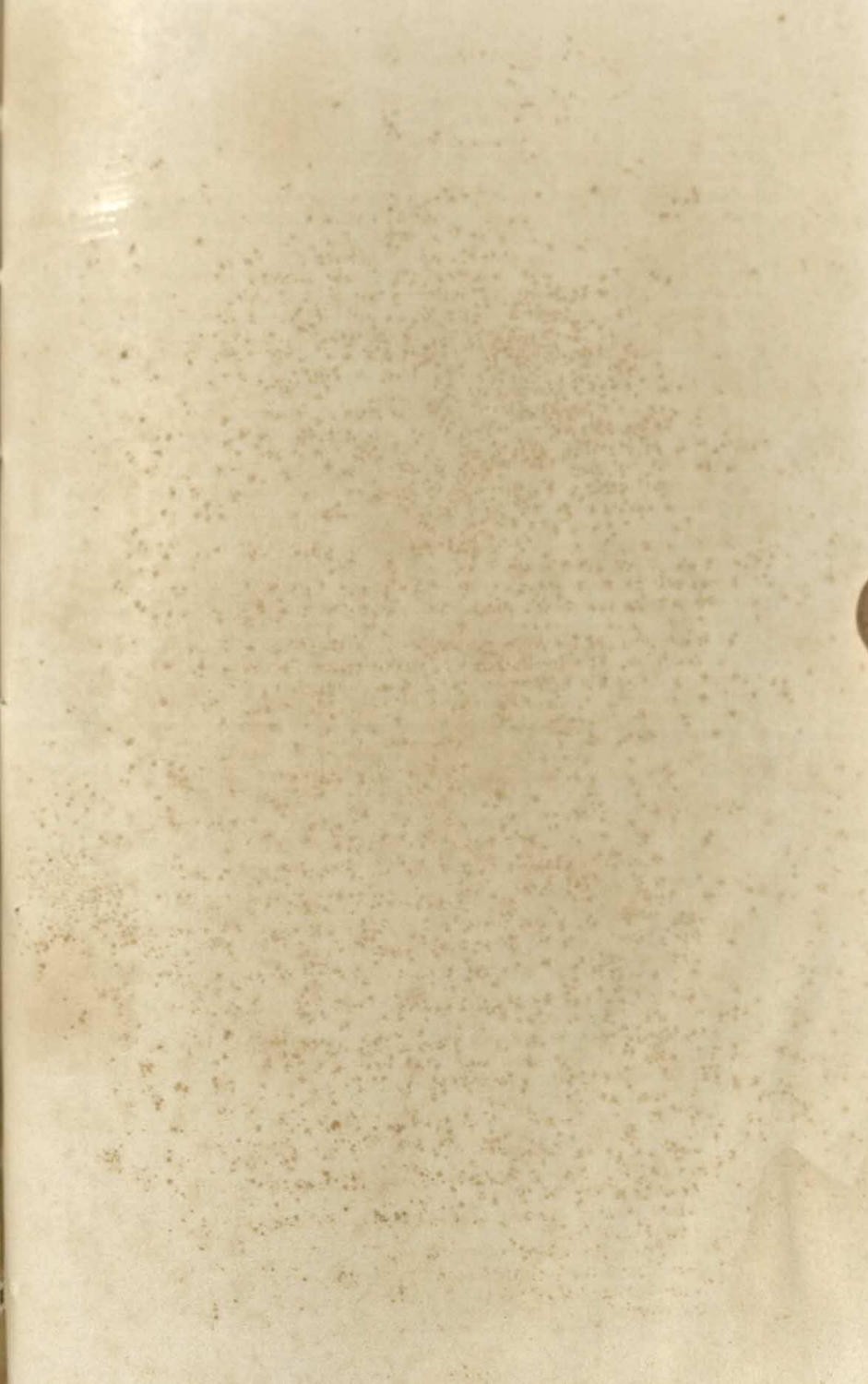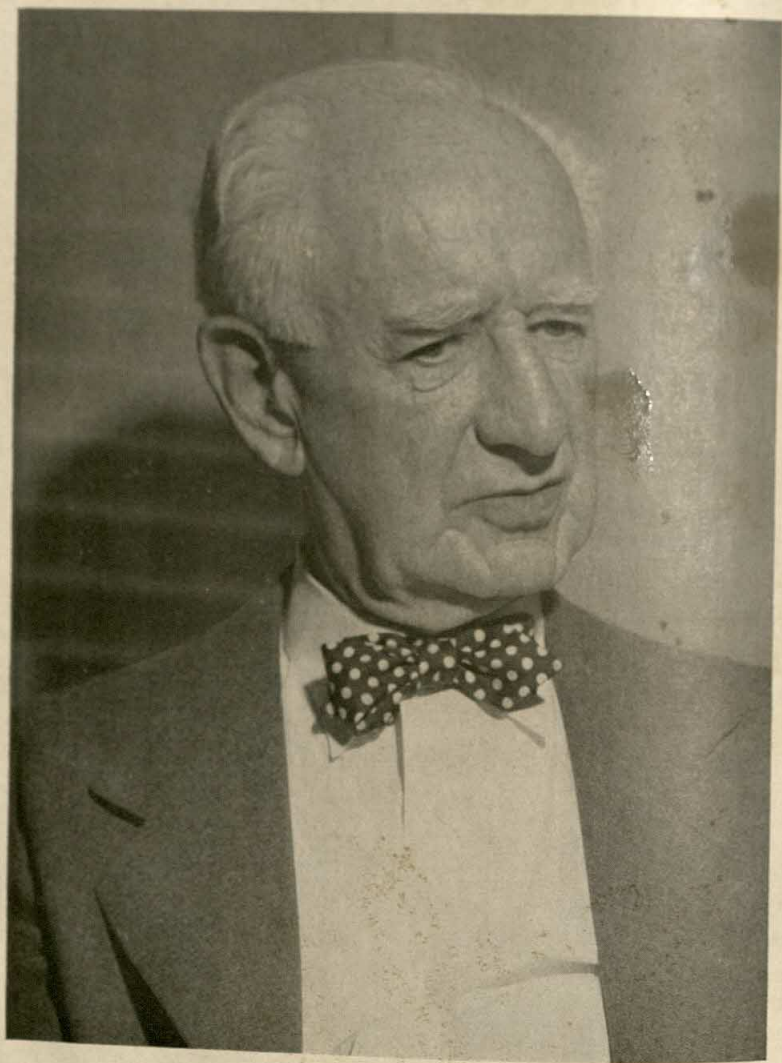
Robert M. Yerkes

THE PERCEPTUAL ELABORATION OF
STROBOSCOPIC PRESENTATIONS

By HANS H. TOCH, Princeton University

The study here reported was designed to explore the visual perception of contiguous stimulus-patterns rapidly succeeding each other in time. In the case of two such consecutive patterns, it has been found that, under given conditions, the first exposed stimulus-figure may not be represented in the configuration that is perceived. The present study was designed to explore further conditions that elicit this response, and to gain qualitative data that could lead to hypotheses about underlying mechanisms.

This phenomenon in question has been recently described by Peter Scheffler.[1] In Scheffler's experiment a horizontal series of five adjacent squares was used. When the second and fourth of these squares were presented simultaneously, and were followed, 0.1 sec. later, by the first, third, and fifth squares, also flashed simultaneously, *only the last three squares were perceived*. Squares 2 and 4 always manifested themselves as dark intervals, especially if 1, 3 and 5 were fixated. Only if Squares 2 and 4 were attentively fixated and full concentration was applied to them, could an indication of light be perceived in them.

Scheffler found the interval of 0.1 sec. between the two exposures optimal for the effect. He also emphasized the importance of similarity of the squares. Lowering or raising the intensity of one of the banks relative to the other, for instance, could offset the phenomenon.

[1] Peter Scheffler, Versuche aus dem Institut für experimentelle Psychologie der Universität Innsbruk: Gestaltete Zeitwahrnehmung (manuscript). Also see Wie sehen wir Bewegung?, *Pyramide*, 10, 1951, 181-184.

Qualitative data showed that the lights in the first bank (2 and 4) were not totally eliminated from the percept. The visible squares looked different when Squares 2 and 4 were omitted from the stimulus-presentation. Small imperfections in the first squares communicated themselves to the three visible squares. On the whole, the nature of observations was subject to considerable variation, and it was even possible to see the three squares flashing twice. Apparent movement was sometimes reported.

The fact that the first of two adjacent and rapidly consecutive stimulus-presentations may not receive proportional representation in the percept has been known for some time. The principal heading under which this phenomenon has appeared in the literature is that of *metacontrast*. Metacontrast has been attributed to an inhibition-effect exercised by the second stimulus, and thus it is defined as "the depressing effect of the second of two brief adjacent asynchronous excitations on the sensory impression produced by the first."[2]

Metacontrast as such was first demonstrated by Stigler in 1910,[3] although precursors have been cited in Sherrington and McDougall.[4] Stigler found that in successively exposing two halves of a geometrical figure (such as a disk) the first half was almost imperceptible to observers, except for its outer border. Later investigators have used presentations in which the second stimulus bordered two sides of the first, and their experiments have determined the conditions optimal for the inhibition of the center flash by the two adjacent flashes following it. The most significant studies in this literature are those of Fry, Piéron, and more recently, Alpern, who has also reviewed earlier work.[5] The findings of these studies may be roughly summarized as follows: (1) The effect increases with the time-interval between 'a' and 'b' flashes, up to 100 m.sec. or 150 m.sec. after which it decreases to zero.[6] (2) The effect increases with the brightness or duration of the 'b' flashes, and decreases with increased brightness or duration of the first flash. (3) The effect takes place only with contiguity or near-contiguity of the flashes.

All the investigators cited have viewed metacontrast basically as a product of neural or chemical inhibition effects in the retina, although Piéron discussed secondary cortical elaborations, such as apparent movements and "eccentricity of rotation

---

[2] Mathew Alpern, Metacontrast: Historical introduction, *Amer. J. Optom.* 29, 1952, 634.

[3] Robert Stigler, Chronophotische Studien über den Umgebungskontrast, *Arch. f. d. ges. Physiol.* 134, 1910, 365-435.

[4] Alpern, *op. cit.*, 631-646.

[5] Glen Fry, Depression of the activity aroused by a flash of light by applying a second flash immediately afterwards to adjacent areas of the retina, *Amer. J. Physiol.*, 108, 1934, 701-707; Mechanisms subserving simultaneous brightness contrast *Amer. J. Optom.*, 25, 1948, 162-178; Henri Piéron, Le processus du métacontraste, *J. psychol. norm. Path.*, 32, 1935, 1-24; Le métacontraste, *Ibid.*, 32, 1935, 651-652; Alpern, Metacontrast, *J. Opt. Soc. Amer.*, 43, 1953, 648-657; The effect of luminance of the contrast inducing flashes on the spatial range of metacontrast, *Amer. J. Optom.*, 31, 1954, 363-369; Alpern, *op. cit., ibid.*, 29, 1952, 631-646.

[6] In discussions of stroboscopic experiments alphabetical order traditionally indicates flashing order. The positions of the letters with respect to each other represents the spatial arrangement of the figures.

of . . . the rings which one does see, which is due to the visual action of the inhibited or unrecognized stimulus."[7]

Research on metacontrast has traditionally been done monocularly. Stigler, however, investigated the possibility of binocular metacontrast using adjacent areas on different retinas. Although he claimed to have been successful he never published relevant data.[8] Baumgardt and Segal, in the course of a paper dealing with the physiology of inhibition and facilitation, also reported having obtained metacontrast across retinae.[9] A confirmation of this finding, needless to say, would make it difficult to think of metacontrast as purely a retinal inhibition.

Werner's *masking experiments* may also be regarded as related to the present study, since the effect obtained was precisely that of the first of two stroboscopically or tachistoscopically presented figures not being perceived.[10] The presentation consisted of two black-on-white figures, the first of which was a solid figure that fitted into the second, a frame for the first. With exposures of 12–15 m.sec., and intervals of 120–240 m.sec., the first figure, according to Werner, "will disappear." Werner regarded such effects as the result of "contour appropriation." According to this hypothesis structuration takes place toward the contour, and to follow up one figure with a frame deprives the first figure of its contour before the completion of structuration, and hence prevents the figure from being built up.[11]

A third type of experiment in which closely similar observations have been made is that of *apparent movement*. Wertheimer made mention of instances in which the first of two stimulus figures was not seen, and movement was observed in the "surviving" figure.[12] His emphasis on "pure phi" suggests the suppression of one or both terminal stages in the percept. More recently, Willey, in the course of experiments concerning combinations of apparent movements, states: "Sometimes there was movement from one member of the first pair of lights to both members of the second pair. *The other member of the first pair of lights was not seen at all,* or was seen as motionless and independent."[13]

A systematic study of apparent movement when alternative terminal stimuli are presented was conducted by Paul von Schiller.[14] His study dealt principally with the factors of time and distance set over against similarity of the initial and final stimuli. Under some circumstances Von Schiller found that the a-figure would be seen as moving to both b-figures, if necessary changing shape en route. This observed tendency of the a-figure to transform itself in every possible respect into the b-figure, Von Schiller called the tendency to total assimilation (*Tendenz zur*

[7] Piéron, *op. cit., J. psychol. norm. Path.*, 32, 1935, 14-15. See also *ibid.*, 24 and 652.

[8] Cited by Alpern, *op. cit., Amer. J. Optom.*, 29, 1952, 631-646.

[9] Ernst Baumgardt and J. Segal, Facilitation et inhibition. Parametres de la fonction visuelle. *Année psychol.*, 43-44, 1942-1943, 54-103.

[10] Heinz Werner, Studies in contour: I. Qualitative analysis, this JOURNAL, 47, 1935, 40-64.

[11] Werner, *op. cit.*, 42.

[12] Max Wertheimer, Experimentelle Studien über das Sehen von Bewegung, *Z. Psychol.*, 61, 1912, 161-265, 217-218.

[13] C. F. Willey Directional variations of apparent movement. *J. exp. Psychol.*, 19, 1946, 656. Italics mine.

[14] Paul von Schiller, Stroboskopische Alternativversuche, *Psychol. Forsch.*, 17, 1933, 180-214.

totalen Angleichung). This is the tendency to become totally figure 'b,' irrespective of its formal constitution. In other words 'b evolves phenomenally from a.'[15]

The basic presentation in the present study consisted of a horizontal series of three adjacent squares, with the middle one flashed before the other two, which could then be flashed simultaneously. This basic presentation was subjected to controlled variations, and was susceptible of expansion through the addition of squares, or the introduction of figural content into squares.

Apparatus. An apparatus was constructed for these purposes that produced light flashes of constant duration and permitted the independent variation of flash intensity and the interval between flashes.[16] Two viewing units allowed for the presentation of figures made up out of 18 squares of light, and also provided facilities for the introduction of detailed figures within the squares.

The apparatus consisted of a timing unit, a set of two switchboards, and the two viewing units. A variac transformer supplied current to the viewing units.

The timing unit was a calibrated mechanical timer, which operated a set of relays. The current for the lights, after passing through the relays, was led to a variac transformer on the control table, the settings of which had been calibrated in terms of brightness of the light sources with a Macbeth illuminometer. The timer itself permitted the variation of the interval between flashes.

The brightness values used in the study ranged from 0.007 to 70 ml. Intervals between flashes, measured from the end of the first flash to the start of the second, ran from an overlap of − 20–200 m.sec. The flash duration was 50 m.sec.

The viewing units consisted of tin-partitioned, black painted compartments, faced by diffusing glass, each compartment containing a 7.5-W filament bulb. The larger unit allowed for the presentation of 2.5 × 2.5 in. squares in any desired pattern. This unit consisted of a honey-comb arrangement of 18 compartments. The smaller unit had three compartments (2.5 × 5 in.), each containing two bulbs and an extra pane of milk glass to increase diffusion. This unit had a space behind its glass front for the insertion of 8 × 10-in. photographic transparencies.

O was seated 7 ft. from the glass screens, on an oculist's chair.

Procedure. Eight *O*s were used for formal study.[17] All were young adults. Each indicated having normal vision, in two cases with corrective lenses. No test for visual acuity was administered.

Sessions lasted 2 hr. and took place in a completely darkened room. Each session was preceded by a 5-min. dark-adaptation period. Observation was free; *O* was simply instructed to report what he saw. The only exceptions to this rule were 'quantitative' sessions in which *O* was asked to note two specific aspects of the perceived configuration. Otherwise, a practice was made to supplement spontaneous

[15] *Ibid.,* 186, 213.

[16] For the design of the apparatus and the wiring of its electrical components, the author is indebted to W. H. Ittelson.

[17] A pilot study preceding formal experimentation was conducted with only one *O*. The investigator wishes to acknowledge, in this connection, the assistance of Edward Engel. One other *O* was tried and excluded from the experiment, since he succeeded in resolving the presentation into three squares under all conditions. The trial data for this *O* are not reported.

reports by questioning about unreported aspects of the experience. Whenever O indicated being unsure about an observation, the presentation was repeated.

Every presentation was run a minimum of 5 times at 2-sec. intervals to permit stabilization of the percept; whenever O was not satisfied that he could make a judgment after 5 presentations, the run was renewed until O reported being sure of what he saw. In the early experiments, each stimulus-condition was presented under a variety of brightness and at least two time-intervals. This practice was later discontinued as being too time consuming. It did not yield additional information. An interval of 40 m.sec. with flash-brightness of 12 ml. was chosen, and all patterns were subsequently presented with these conditions constant. Every pattern was repeated at least twice. If the two responses were not identical, the experiment was repeated until two consecutive reports were judged to be the same.

Fixation. The data were collected under conditions of free viewing. O was instructed to look at the presentation, and he was left free to attend to any of its aspects. Since fixation was omitted as an experimental variable, some control data were taken on two types of fixation.

The first type involved fixation of the center of the presentation (the locus of the 'a' square) without, however, a fixation-stimulus.[18] Incidental findings with such fixation confirmed Scheffler's observation that, under conditions other than optimal, fixation of the center space facilitated the perception of light there.

A case illustrating the effect of central fixation was that of a naïve O who in his first session never reported light between the two squares. In subsequent sessions, he began seeing light under some conditions, a fact he attributed to having "switched" to fixating the center of the presentation. It was his impression that he had initially "followed the two moving squares," and thus prevented himself from seeing anything between them.

Conversely, the protocols of the one O who was discarded because he reported the center square throughout, indicate that his performance was probably due to an analytic set, and an unusually careful fixation.[19]

A second type of fixation investigated was peripheral fixation. As a check on this, a fixation-lamp was constructed which could be placed anywhere in the experimental room. Three Os were used for one session each with this fixation-light.

The light was placed some distance from the viewing unit, above it, below it, or to one side. O was instructed to fixate, and to watch the presentation without abandoning fixation. In all cases, and with all the presentations used, this results in the Os perceiving no center squares, but *not seeing any movement*. The squares were reported as simply "flashing on."[20]

---

[18] The use of a fixation-element is precluded by the fact that it would probably alter the nature of the presentation.

[19] It has been noted that 'analytic' set is unfavorable to apparent movement-perception also. Von Schiller, *op. cit.,* 180; G. M. Stratton, The psychology of change: How is the perception of movement related to that of succession?, *Psych. Rev.* 18, 1911, 262-293.

[20] This may be of relevance to apparent movement. As may be recalled, in apparent movement-situations, fixation to one side of a presentation favors movement to the fixated side. See H. G. Van der Waals and C. O. Roelofs, Optische Scheinbewegung, *Z. Psychol.,* 114, 1930, 241-288; 115, 1930, 91-190; Von Schiller, *loc. cit.*; Wertheimer, *loc. cit.*

EXPERIMENTAL CONDITIONS AND RESULT

The 30 experimental conditions of the present study can be summarized under 3 headings.[21] (*a*) The first seven may be characterized as *basic patterns*. They were designed to explore the important features of the three-square arrangement. (*b*) The second heading comprised ten *figure-ground* conditions. They were so designated because the figural content was systematically introduced into the three-square presentation, which thus became ground. (*c*) The remaining conditions involved the formation of a sequence. Presentations of more than three squares were used here, in an effort to explore their differential perceptual elaboration.

(*a*) *Basic Patterns.* The procedure used in the first of these experiments was, Condition 1, to present the figure with instructions to "describe it in detail," and, Condition 2, to flash the figure on the same screen with a set of two 'b' squares alone, *i.e.* with an actual space between them. *O* was asked to specify differences, if any, between the two perceived configurations in Condition 1.

Three *O*s were used for the purpose of more closely specifying the effects of varying light-intensity and time-interval between exposures. They were presented with Condition 3, the figure over eight brightness-values (ranging from 0.007 to 70 ml.) and 11 time-intervals ($-$ 10–160 m.sec.)[22] They were instructed to report (a) what, if anything, they saw in the center of the presentation; and (b) any movement they perceived. The whole series was carefully repeated with one *O*, both binocularly and monocularly, to determine intra-*O* reliability and monocular-binocular differences, if any. Intervals were used in random order, and intensities were randomized within intervals.

A Polaroid viewing unit was constructed with filters aligned so that the center square would be seen by one eye, and the outside squares by the other. This is called Condition 4, and was used to test the central as opposed to the peripheral locus of the intraction.

The effects of contiguity and figural similarity were explored with Conditions 5 to 7, which are shown in Fig. 1.

---

[21] Complete presentation of the full set of experimental conditions, and of the results obtained with each, may be found in Hans Toch, Perceptual elaboration of a stroboscopic presentation of three contiguous squares, Doctoral dissertation, Princeton Univ., 1955.

[22] The brightness-values used were .007, .05, .4, 5, 12, 19, 45 and 70 millilamberts. The intervals were $-$ 10, 0, 10, 20, 40, 70, 100, 120, 130, 140, and 160 milliseconds.

The results were as follows:

Results. (1) The basic (three-square) figure was perceived by all *Os* as two squares, in most cases "moving away from the center," and otherwise "restless and vibrating."

(2) When compared with the set of 'b' squares in Condition 2, the
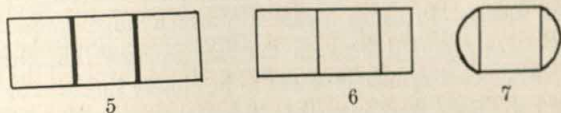
FIG. 1. GRAPHIC SUMMARY OF BASIC CONDITIONS 5-7

three-square presentation was found in all cases to differ in that it moved, "while the other squares were steady," and it was noticeably brighter than the comparison squares.

(3) The quantitative data for the one *O* with whom the series was repeated are plotted in Fig. 2. These data, and those obtained from the two other *Os*, can be summarized as indicating (a) that increasing brightness favored the perception of only two squares and the three squares tended
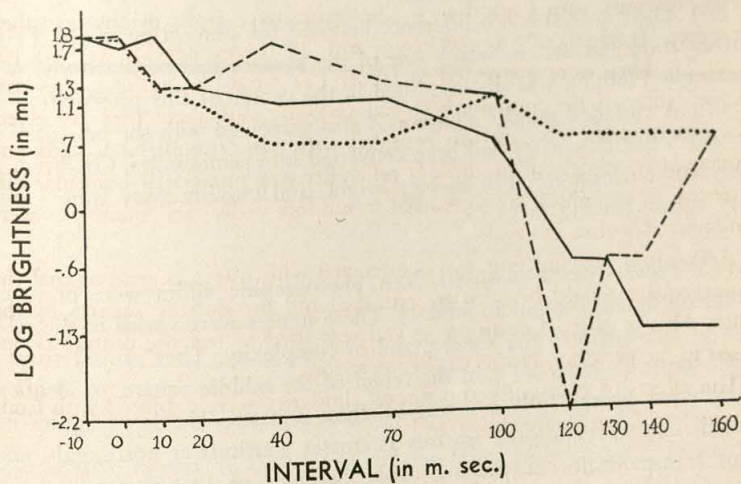
FIG. 2. PERCEPTION OF TWO SQUARES AROUND EMPTY SPACE AS A FUNCTION OF BRIGHTNESS OF FLASH AND INTERVAL BETWEEN FLASHES

Data for one *O*, three sessions, and one monocular and two binocular. The curves represent lower thresholds for 'complete disappearance.' At values below each curve light was perceived between the two squares. At values above the curves the space between squares was reported as empty. The solid line and the broken line represent the first and second binocular observations respectively. The dotted line represents monocular data.

to be perceived only at the lower intensities; and that (b) the brightness range within which only two squares were perceived increased with the interval up to the point where the effect was no longer operative. Reports of "some light" or "flash in the center" were obtained for intermediate conditions, between those favoring a two-square as against a three-square resolution. The perception of apparent movement was directly related to that of two squares. In every instance in which no movement was perceived, light was reported between the two squares. The thresholds of these two phenomenal results appeared to be closely linked. For the one O with whom conditions were repeated, observations proved highly reliable. In monocular observations the phenomenon appeared to be less favored by longer intervals than in binocular ones.

(4) Viewing through the Polaroid eye-piece did not result in reports differing from observations without it. In all cases, two squares were seen moving outward.

(5) Contiguity of the squares proved to be essential to the perception of only two squares. A ¼-in. strip between squares, Condition 5, enabled most Os to perceive a middle square, which was seen as dimmer or of shorter duration than its neighbors.

(6) There was also no difference between the perception of the 'basic' presentation and that in which one of the squares had been reduced to a rectangle. Nothing was distinguished in the center of this presentation.

(7) A curtailed middle square was also perceived with the presentation in which the 'b' squares had been converted into semi-circles, Condition 7. The semi-circles were seen as very bright, and moving away from a dim grey square linking them.

(b) *Figure-ground patterns.* Ten photographic slides were prepared featuring figural content in squares. These figures were varied in the relation of the figure to the square, and in complexity. They ranged from a small figure (a letter 'H') in the center of the middle square, to identical or different figures fully covering all three squares. The large figures included lines traversing the squares as crosses, verticals or horizontals, and color transparencies of playing cards. Samples of the figures used are shown in Fig. 3 and Fig. 4.

Results. Out of the 10 slides only 2 resulted in the perception of nothing in the center of the presentation. These are shown in Fig. 3. One of these featured horizontals running through the three squares, Condition 14; the other consisted of three identical playing cards, the queen of hearts. In all other cases, including a slide featuring vertical lines on all three squares,

Condition 15, and one containing playing cards in which the center card (a queen of hearts) was dissimilar from the other two (backs of playing cards), *the figures of the center square were perceived.* In the case of all
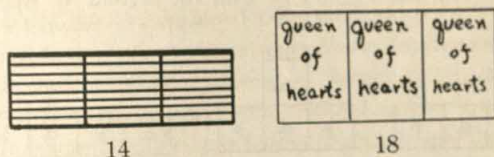
FIG. 3. CONDITIONS 14 AND 18

line drawings other than Condition 14, the figures were reported on a dark ground surrounded by a faint halo of light, just sufficient to make them perceptible.[23] Two squares were in each case seen to move away from
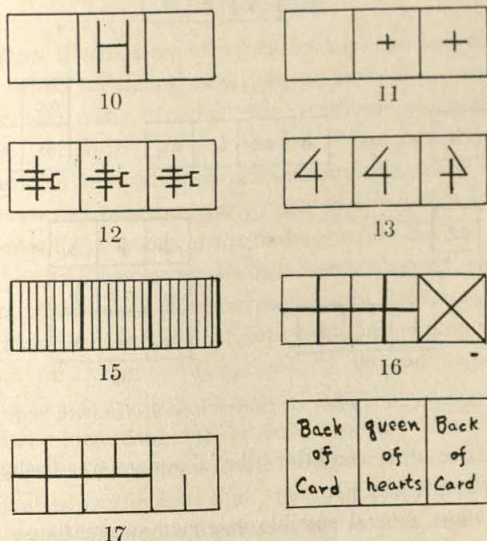
FIG. 4. CONDITIONS FEATURING RESIDUAL FIGURAL CONTENT

these residual figures. In the case of the dissimilar playing cards, the entire center card was perceived.

(*c*) *Conditions involving sequence.* In an attempt to investigate the

[23] The above was the case with binocular observation. Monocularly, there was a tendency not to perceive anything in the center. This fact may be accounted for by the residual figure's being monocularly mislocalized in these cases, and transferred, in apparent direction, into one of the 'b' squares in which it then appeared.

perceptual results of expanding the basic presentation, the latter was made a part of eleven different complex presentational arrangements. In each case, one or more squares were added, either in phase with the center 'a' square (20, 21 and 22, Fig. 5) or with the second 'b' squares (23 and
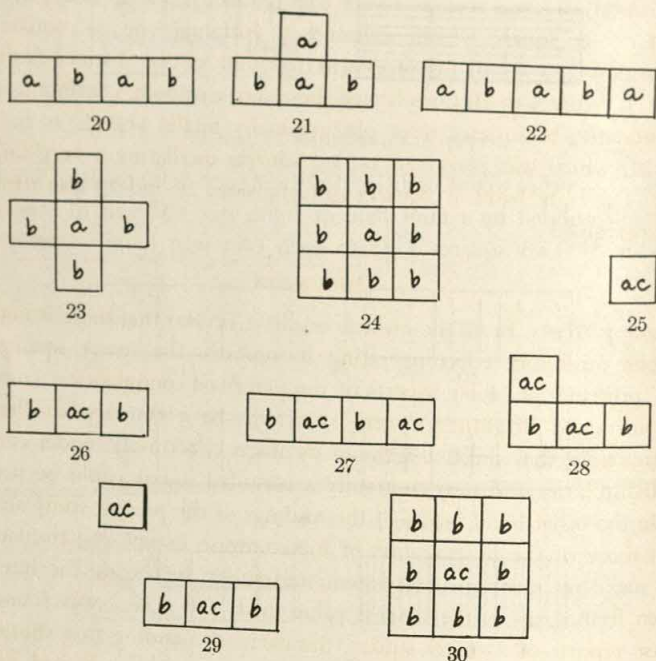
FIG. 5. GRAPHIC SUMMARY OF CONDITIONS INVOLVING SEQUENCE

24, Fig. 5). In five other conditions the 'a' square was flashed again after the 'b' squares (26-30, Fig. 5).

Results. The most general possible statement summarizing findings with these presentations is that the 'a' square was perceived in a variety of ways, depending on the nature of the sequence. Thus, (1) unidirectional arrangements (such as Condition 20) resulted in smooth unidirectional movement, with the 'a' square as a movement-terminal;[24] (2) centripetal sequences (such as Condition 22) tended to make the 'a' square perceived as a square; (3) total enclosure of the 'a' square (Conditions 23, 24) resulted, as a rule, in its perception as a flash of light; (4) in presentations

---

[24] Movement-terminals are the end-points of a movement percept, *i.e.*, the starting place and point of arrival of the moving object which is perceived.

in which the 'a' square was flashed again as a third bank, the 'b' squares were perceived as phases linking 'a' to 'c.' Thus, the 'b, ac, b' presentation (Condition 26) was most often reported as an outward-inward movement (predominantly inward) of dim light, which 'formed' a very bright center square. Similarly, an 'ac' square totally surrounded by (six) 'b' squares (Condition 30) was seen as a very dim reddish frame around a brilliantly bright center square which outlasted it. Although the 'ac' square alone (Condition 25) in all instances was reported as two distinct flashes, adjacent 'b' squares in all cases linked these into a smooth, continuous square. Occasionally 'b' squares were phenomenally totally absent, as in Condition 27, which was perceived as two squares oscillating a very small distance. The right hand 'b' squares in Conditions 28 and 29 also tended not to be perceived.

DISCUSSION

Basic patterns. From the present results it is clear that there is not an all-or-none inhibitory effect operating to suppress the center square in the 'bab' presentation. First, aspects of the perceived configuration (movement and increased brightness) were shown to be a function of the center square, since they did not occur in its absence. Secondly, under conditions of dissimilarity and near-contiguity a curtailed square could be perceived.

On the other hand, many of the findings of the present study are in line with those of the investigators of metacontrast, except that the latter usually make no mention of apparent movement. Increasing the interval between flashes, up to the longest value used, 160 m.sec., was found to decrease reports of a center square. Similarly, the finding that shorter intervals are effective only with increased brightness is in accord with data on metacontrast regarding the brightness of the second bank of lights. Given the fact that the second bank of lights (with double the area of the 'a' square) has twice the total light energy of the first, an increase in total brightness would preferentially favor the second bank. Another finding with regard to metacontrast, that of the vital role of contiguity, was confirmed.

In the present experiment it was found that the effect was not strikingly diminished when the first and second exposure were to opposite eyes. This is in line with findings concerning flicker fusion and apparent movement, both of which have been obtained across the two retinas.[25]

---

[25] Among the studies reporting flicker fusion across the two eyes, see F. Schwartz, Ueber die binokulare Summation von Flimmerlicht, *Z. Sinnesphysiol.*, 70, 1943,

Figure-Ground conditions. The perception of residual figures on barely perceptible ground in the empty space corresponding to the 'a' square, suggests that prior to elaboration of the figure it must have been divorced from the square that contained it, which was not itself elaborated. It becomes possible to think of some sort of classification or sorting process preceding actual perceptual structuration.

The question may be posed as to why horizontals extending throughout the presentation (Condition 14), and the three identical picture cards (Condition 18), are not subject to the same two-category sorting as other presentations. Why, in other words, is all of 'a' in these two cases unelaborated, in view of the fact that content is elaborated in comparable presentations?

In the case of the horizontals, it may be speculated that these are not perceived because they are perceptually incorporated into the lines which *are* perceived (those in the 'b' squares) of which they are in fact continuations. Similarly, it may be assumed that the identity of the three queens permits their perceptual identification, and hence the incorporation of the 'a' queen.

The fact that a whole queen is perceived when followed by two backs permits the hypothesis that the 'figure-ground phenomenon' in line figures is attributable to some sort of conceptual separability of figure and ground in these presentations.

Spatio-temporal sequence. It has recently been suggested that "the facts of (movement) experiments can be explained by the hypothesis that the retina responds to adjacent and successive order. . . . The stimulus for motion . . . may be *ordinal.*"[26] The indications provided by the 'sequence' conditions of the present study are that, as a descriptive statement, this may hold true of more than movement perception; characteristics of configurations perceived with rapid successive stimulation appear to be intimately linked to the ordinal aspect of the stimulus-situations. Qualities in the percept such as apparent brightness and perceived duration, for instance, are subject to variation by means of differential spatio-temporal arrangement in the presentation.

22-29. G. J. Thomas has reviewed relevant literature in his study: The effects on critical flicker frequencies of interocular differences in intensity and in phase relations of flashes of light, cff, this JOURNAL, 67, 1954, 623f. With regard to apparent movement see K. R. Smith, Visual apparent movement in the absence of neural interaction, this JOURNAL, 61, 1948, 73-78, and H. S. Langfeld, Apparent visual movement with a stationary stimulus, this JOURNAL, 39, 1927, 343-355.
    [26] J. J. Gibson, The visual perception of objective motion and subjective movement, *Psychol. Rev.*, 61, 1954, 310.

What are the dynamics of this relationship? It has been suggested, in another context, that movement perception results when stimulus-situations are interpreted in line with an unconscious generalization derived from early experiences with moving objects.[27] The common denominator in all such experiences (and hence, the basis of the generalization) would be the successive stimulation of adjacent portions of receptor surfaces. More specific situational denominators are, however, involved.

Object connotation and sequence. A moving object manifests itself as a unified experience. The movement of the object cannot be divorced from its other connotations in the percept. That this unity goes deeper is exemplified by the demonstration of the rotating trapezoidal window. The perception of this device as a rectangle in perspective involves seeing it in oscillatory movement, whereas its correct identification (as a trapezoid) shows it as rotating.[28] The stimulus-situation is here interpreted by the perceiver, *in toto,* and all the attributes of the percept arise from the meanings assigned to the whole.

The elaboration of a movement-configuration does not consist of separate processes making for 'object' and 'movement,' but of a single process culminating in the perception of an object-in-movement. The sequential aspects of this configuration are inseparable from its meaning aspects.

Thus, when the 'bab' presentation is reported seen as "two squares moving away from the center," the question "where is the third square?" does not arise. The configuration called for by the ordinal aspects of the stimulus condition is a centripetal one, culminating in two objects. Object characteristics are assigned to the presentation to achieve this phenomenal result. A center square would not make sense in this context, and hence is not elaborated into the configuration.

Classification and disposition in sequences. If configurations are structured in line with meanings assigned to presentations, it is important to inquire in what manner and on what basis meanings are assigned.

To formulate these questions, they have to be divorced from each other. Accordingly, perceptual elaboration may be conceived of as a two-stage process. The first stage would consist of *classification*[29] of the stimulus-condition (the assignment of meaning to it), and the second of the *disposition* of relevant perceptual data into a configuration.[30] By means of this distinction one avoids the confusion of product with determinants of processing which appears to underlie much of structuralist theorizing.

It would be the purpose of the perceptual classification process to determine the nature of configurations, by means of a rapid unconscious evaluation or physiological categorizing of perceptual data. Actual disposition would follow accordingly.

[27] H. H. Toch and W. H. Ittelson, The role of past experience in apparent movement, *Brit. J. Psychol.,* (publication pending).

[28] Adelbert Ames, Jr., Visual perception and the rotating trapezoidal window, *Psychol. Monogr.* 65, 1951, (No. 324), 1-32; F. P. Kilpatrick and W. H. Ittelson, Three demonstrations involving the visual perception of movement, *J. exp. Psychol.* 42, 1951, 394-402.

[29] The author wishes to acknowledge his indebtedness to Adelbert Ames, Jr., for the suggestion of this term.

[30] Processes being indivisible in fact, the order in which the above has been conceptualized is, of course, arbitrary. Classification would have to take place in terms of the disposition called for.

In conditions involving sequence, the nature of the classification clearly determines the character of the disposition. Identical physical conditions (classified differentially because they occur in different presentational sequences) may be elaborated as bright squares, dim squares, brief squares, squares of longer duration, amorphous flashes of light, movement terminals, vague feelings "that something might be there," or as nothing at all phenomenally isolable.

The concrete disposition is a function of the role assigned to impinging events in whatever sequence is meaningful to the perceiver. The overall attempt is to perceptually create a world which behaves reliably, despite constant and continuous flux.[31]

---

[31] This same point has been made by Ittelson in The constancies in perceptual theory, *Psychol. Rev.*, 58, 1951, 285-294.

EXTINCTION FOLLOWING PARTIAL REINFORCEMENT WITH CONTROL OF STIMULUS-GENERALIZATION AND SECONDARY REINFORCEMENT

By D. W. Tyler, University of Texas

Reinforcement theorists have attempted to deal with the problem of partial reinforcement in terms of 'stimulus-generalization' and 'secondary reinforcement,' and their interpretation has received some support from experiments by Sheffield[1] and Rubin.[2] Sheffield distributed practice to control stimulus-generalization and found no difference in the resistance to extinction of a running response in partially and consistently reinforced animals. Rubin distributed practice and manipulated goal-box characteristics to control both stimulus-generalization and secondary reinforcement, and he found superior resistance to extinction of a panel-pushing response in consistently reinforced as compared with partially reinforced animals.

Recent experiments from this laboratory have supported an alternative interpretation, which has been called the *discrimination-hypothesis*,[3] and which is based on the assumption that the animal may learn a great deal about the characteristics of the training situation—about the occurrence of reinforcement and non-reinforcement, about simple sequences of reinforcement and non-reinforcement, and about stimuli associated with reinforcement and non-reinforcement. The hypothesis asserts that rate of extinction is a function of the readiness with which the animal is able to discriminate, in terms of these learned characteristics, the transition from

training to extinction. From this point of view, a number of experimental results have been predicted, which cannot be dealt with in terms of stimulus-generalization and secondary reinforcement. Controlling stimulus-generalization by substituting delayed reinforcement for non-reinforcement in the partial training procedure does not eliminate the difference in rate of extinction between partially and consistently rewarded groups.[4] Random partial reinforcement results in greater resistance to extinction than does regularly alternating partial reinforcement despite the fact that precisely the opposite result is predicted from the principle of stimulus-generalization.[5] Finally, the discrimination-hypothesis yields the experimentally confirmed deduction that under certain conditions extinction will be *more* rapid when response is followed by stimuli previously associated with reinforcement than when it is followed by stimuli previously associated with non-reinforcement.[6] While these results do not fit the reinforcement theory, the results of Sheffield and Rubin do not fit the discrimination-hypothesis. It may be well, therefore, to reëxamine the latter findings, especially since the design of Rubin's experiment is so cumbersome, and since information on the outcomes of both experiments is so limited. Rubin presents no data on latency of response during extinction, although acquisition was measured in terms of latency, while Sheffield's data on running time are presented in an unusual and probably quite insensitive form.

Problem

In the present study, the essential features of the experiments by Sheffield and Rubin were reproduced with three groups of rats. All groups, hunger-motivated, learned to traverse a runway and jump to a goal-box under conditions of distributed practice. Group *Cn* was rewarded on every training trial in a goal-box of a given color (*e.g.* white). Groups *Sa* and *Rv* were randomly rewarded on only 50% of the training trials. The goal-box for Group *Sa* was the same on reinforced trials and non-reinforced trials (*e.g.* white), while Group *Rv* was rewarded in a box of one color (*e.g.* white) and non-reinforced in a box of the opposite color (*i.e.* black). On extinction trials (also distributed) all groups found the goal-box previously associated with reinforcement. The three groups

[4] Crum, Brown, and Bitterman, *op. cit.*, 228-237.
[5] Longenecker, Krauskopf, and Bitterman, *op. cit.*, 580-587; Tyler, Wortz, and Bitterman, *op. cit.*, 57-65.
[6] Bitterman, Feddersen, and Tyler, *op. cit.*, 456-464; Elam, Tyler, and Bitterman, *op. cit.*, 381-384.

duplicate in all essential respects the crucial conditions of both Sheffield and Rubin. Group *Cn* and Group *Sa* permit the critical comparison made by Sheffield (consistently versus partially reinforced groups with distributed practice to control for stimulus-generalization) while Group *Cn* and Group *Rv* permit the comparison made by Rubin (like that of Sheffield with control for secondary reinforcement as well as for stimulus-generalization).

Since training trials are distributed in all three groups (15 min.), the carryover of afferent components from trial to trial should be negligible and none of the groups should be favored from the standpoint of stimulus-generalization. According to reinforcement theory, therefore, the prediction of differences in rate of extinction among the three groups must be based upon the amounts of reinforcement (both primary and secondary) given each group. Group *Rv* receives the same number of primary reinforcements as Group *Sa,* but less secondary reinforcement on non-reinforced training trials. Group *Rv* should consequently extinguish more rapidly than Group *Sa,* and more rapidly than Group *Cn* which is primarily reinforced on every trial. Whether Group *Sa* extinguishes more rapidly or at the same rate as Group *Cn* depends on the relative strength of primary and secondary reinforcement and on the rate at approach of the two growth-functions to their asymptotes, but under no circumstances can Group *Sa* be expected to extinguish less rapidly than Group *Cn.*

The discrimination-hypothesis yields a quite different prediction with regard to rate of extinction in the three groups. On the basis of the dissimilarity of stimulating conditions encountered during training and extinction, Group *Cn* should extinguish *most* rapidly. Ss of this group encounter non-reinforcement for the first time during extinction and in a goal-box previously associated with reward. They should consequently have little difficulty in discriminating between the two series of experiences, and extinction should be very rapid. Groups *Rv* and *Sa* both encounter non-reinforcement during training and for them the transition from training to extinction is less abrupt and therefore less discriminable. The transition should be least abrupt for Group *Sa,* which during training experiences non-reinforcement in the goal-box to be encountered in extinction, and it should probably extinguish less rapidly than Group *Rv,* which is non-reinforced for the first time during extinction in the goal-box consistently associated with food during training. Thus, while reinforcement theory predicts the order of extinction as *Rv—Sa-Cn,* the discrimination-hypothesis predicts *Cn—Rv-Sa,* with the dash indicating a larger difference than the hyphen.

METHOD

Subjects. Thirty-six experimentally naïve Albino rats were studied. They ranged in age from 5 to 6 mo. at the beginning of the experiment.

Apparatus. The apparatus employed consisted of a straight, elevated runway and a series of interchangeable goal-boxes. The runway was 3.75 in. wide, 7.66 ft. long, and 9 in. from the end was a separate stand which held a goal-box to which the animal was required to jump after traversing the runway. The goal-box had a 6×6 in. window containing an unfastened black-and-white vertically striped card which served as a target and which fell backward to admit the animal. It prevented *S* from seeing the interior of the goal-box until the terminal response was made. A plywood funnel-arrangement fixed to the window of the goal-box served to direct the jumping-response to the card. The entire apparatus was painted mid-gray with the exception of the goal-box *interiors* which were either black, white, or mid-gray.

Procedure. Two weeks prior to preliminary training, all animals were placed on a 24-hr. feeding schedule. During this time, each *S* was gradually reduced to 80% of its satiated bodily weight. On the basis of the amount of weight-loss in the previous 24-hr. period, it was possible to feed each animal enough wet mash to bring its weight to the 80% level at the beginning of the following day's trials. This feeding procedure was employed throughout the experiment.

Preliminary training consisted in teaching the animals to jump gradually increasing distances from the end of the runway to a gray goal-box (gray inside and out). They were first trained to jump to the open window and then to the unfastened striped card. At the end of one week, all animals were readily jumping the 9-in. gap to the card in the goal-box window. Each jump during this period was rewarded by a few seconds of feeding with wet mash inside the gray box.

Experimental training. At the end of the preliminary training period, the animals were divided into three groups equated for jumping ability and general adjustment to the situation. All animals were then given 10 trials per day for 14 days. On each trial, an animal was required to traverse the length of the runway and jump the 9-in. gap to the goal-box. Upon entering the box, the animal was confined there for a period of 20 sec. whether the trial was reinforced or non-reinforced. At the end of this interval, it was removed and placed in a mesh waiting cage until the start of the next trial. The time required for an animal to traverse the runway and jump to the goal-box was recorded by means of a stopwatch. Any animal which failed to reach the goal-box within 90 sec. was manually guided by *E* until it did so. The training schedule was so designed that the time between successive trials for any given animal was never less than 15 min., and often more.

For one group of Ss, Group *Cn,* reinforcement was given on each training trial. Half the animals of this group were fed in a black box (black interior) while the other half were fed in a white box (white interior). The remaining two groups, like Group *Cn,* were given 10 trials per day but were reinforced on only 50% of them. Half the animals in each group were fed in a white box and half in a black box. Training differed for these two groups only on those trials which were *not* reinforced. In one of the partially reinforced groups, Group *Sa,* animals which were reinforced in a black box also found a black box on non-reinforced trials, while animals reinforced in white found a white box on non-reinforced trials. In the second partially reinforced group, Group *Rv,* animals rewarded in white en-

countered a black box on non-reinforced trials and those reinforced in black en-
countered a white box on the non-reinforced trials. Reinforcement and non-rein-
forcement were administered according to selected Gellermann orders in both
groups.[1] To insure the absence of food-particles on non-reinforced trials, inter-
changeable goal-boxes, two white and two black, were employed. One box of each
color was used for reinforced trials, and the other was used only for non-reinforced
trials.

Extinction. During extinction, the animals were given 10 non-reinforced trials
per day. The inter-trial interval during this phase was again never less than 15
min. for any animal. Treatment of the three groups was identical during this por-
tion of the experiment. All animals encountered a goal-box of the same color as
that previously associated with food; that is, animals previously rewarded in a
white box now encountered a white box without food, and those rewarded in a
black box encountered a black box without food. Animals were confined in the empty
goal-box for 20 sec. on each trial, after which they were removed to the mesh
waiting cage where they remained until the start of the next trial. If at any time
during extinction an animal failed to traverse the runway and jump to the goal-
box within 90 sec., it was removed from the runway and placed in the mesh wait-
ing cage where it remained until time for its next trial. The criterion of extinction
was failure to make the appropriate response within 90 sec. on two consecutive
trials.

RESULTS

In Fig. 1, the performance of the three groups during the 14-day
training period and the 3-day extinction period is plotted in terms of
mean log time per trial. Table I summarizes the performance of the

TABLE I

MEAN LOG TIME PER TRIAL DURING TRAINING AND EXTINCTION FOR
EACH OF THE THREE GROUPS

Group	Training	Diff. from Cn	Extinction	Diff. from Cn
Cn	1.3701	—	1.7120	—
Rv	1.4605	−0.0904	1.5101	0.2019*
Sa	1.3806	−0.0105	1.4563	0.2557†

* Significant beyond the 5% level of confidence ($t = 2.80$).
† Significant beyond the 1% level of confidence ($t = 3.15$).

groups during both phases of the experiment in terms of the same measure.
The course of learning was quite similar for all three groups, and the
slight differences among them did not attain significance at any time during
the training period. It appears, therefore, that neither the additional pri-
mary reinforcement afforded the animals of Group Cn relative to the other
two groups, nor the greater secondary reinforcement afforded Group Sa
relative to Group Rv, had any measurable effect on the rate of acquisition.

[1] L. W. Gellermann, Chance order for alternating stimuli in visual discrimination
experiments, *J. genet. Psychol.*, 42, 1933, 356-360.

While the learning curves are strikingly similar, the three extinction curves are strikingly dissimilar. The two 50% groups show a moderate increase in running time from the 14th training day to the 1st extinction day, whereas the 100% group is responding as slowly after only 10 extinction-trials as when training was begun. The difference between the 100% group and each of the 50% groups is significant beyond the 1% level
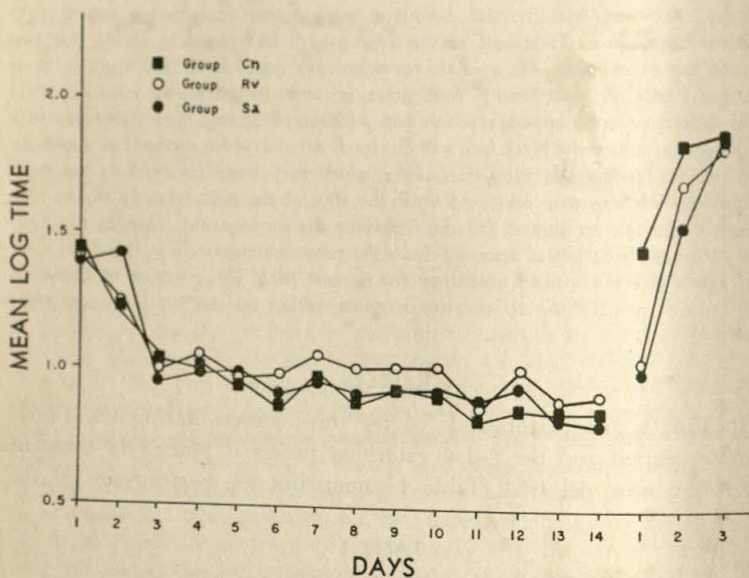
FIG. 1. MEAN LOG TIME PER TRIAL PER DAY FOR THE THREE GROUPS DURING TRAINING AND EXTINCTION

($t = 4.26$ and 4.65) for the first day of extinction, although the two 50% groups do not differ significantly on the first day.

A somewhat more detailed picture of the course of extinction is presented in Fig. 2, which shows the mean log time of each of the three groups on each of the last 10 training trials and each of the 30 extinction trials. There it will be noted that the difference between the 100% group and the two 50% groups, although initially absent, develops rapidly during the first 10 extinction trials. The running times of the 100% animals are significantly longer than those of the other two groups by the end of this period, although the two 50% curves are almost identical. By the 12th trial, a difference between the two 50% groups is discernible—the animals of Group Rv tending to extinguish more rapidly than those of Group Sa. The mean log time for Group Sa on the second day of extinction is

1.53 sec., while the time for Group *Rv* is 1.68 sec. The difference, however, falls short of the 5% level of significance ($t = 1.27$). Considering all three extinction days together, the difference between Groups *Cn* and *Sa* is significant beyond the 1% level ($t = 3.15$), while the dif-
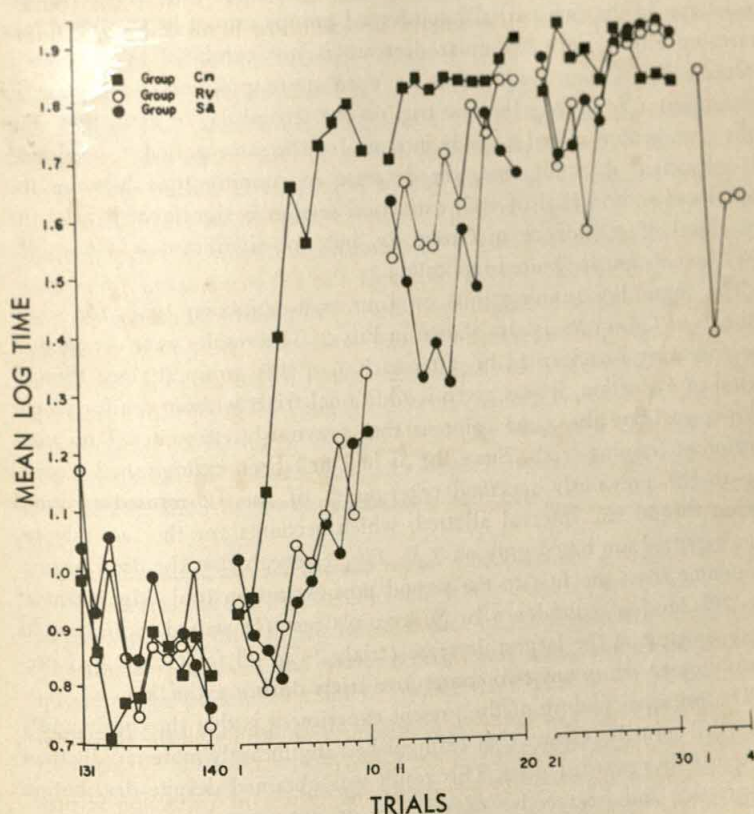
FIG. 2. MEAN LOG TIME PER TRIAL ON THE LAST DAY OF TRAINING AND ON THE THREE DAYS OF EXTINCTION FOR ALL GROUPS, AND ON A POST-EXTINCTION DAY FOR GROUP *Rv*

ference between Groups *Cn* and *Rv* is significant beyond the 5% level ($t = 2.80$). The differences between the two 50% groups is not statistically significant ($t = 0.59$).

Especially interesting are the running times for the three groups on extinction trials 1–4 and 11–15. On the first extinction trial, it will be noted, all three groups are at essentially the same level. On the second and third trials, running times increase in the 100% group and decrease

in the two 50% groups. The only real difference in the two partially reinforced groups during this period is the point at which the curves begin to ascend. This point is Trial 4 for Group *Rv*, but is delayed until Trial 5 for Group *Sa*. Meanwhile, the curve of the 100% animals is rapidly approaching a maximum. The decrease in running time following the initial run in the two partially reinforced groups cannot be attributed to a warm-up effect, since a similar decrease is not exhibited by the 100% animals. For Group *Sa* there is an even more pronounced decrease in running time following the first trial on the second day of extinction. The other two groups reveal a steady increase for the same period. Considering all extinction days together, the decrease in running time between the initial and second trial of each extinction session is significant beyond the 5% level of confidence in Group *Sa*, but not significant in Group *Rv* (Wilcoxon's test for paired replicates).[8]

The mean log running time on four post-extinction trials for seven animals of Group *Rv* is also shown in Fig. 2. The results were obtained in the following manner: 24 hr. after each *S* in this group attained the criterion of extinction, it was given 4 additional trials without reinforcement with a goal-box the same color as that previously encountered on *nonreinforced* training trials. Since the *S*s had first been extinguished according to the previously specified criterion, 5 of the 12 refused to jump within the 90 sec. interval allotted, which accounts for the fact that results reported are based only on 7 *S*s. Fig. 2 reveals that the drop in running time from the first to the second post-extinction trial (significant at the 2% level of confidence by Wilcoxon's test for paired replicates) is twice as great as the largest decrease (trials 23 to 24 for Group *Rv*) previously observed on any two consecutive trials during extinction.

The principal finding of the present experiment is that the continuously rewarded animals (Group *Cn*) extinguished significantly more rapidly than the partially rewarded ones. This result was obtained despite distribution of practice and irrespective of whether or not secondary reinforcement was minimized on non-reinforced training trials for the partially reinforced animals. This finding is contrary to results reported previously by both Sheffield and Rubin. (Sheffield compared rates of extinction of a running response in two groups of rats trained and extinguished in a manner analogous to the present Groups *Cn* and *Sa*, and found no statistically significant difference between them, but a difference did appear in the present experiment. Rubin, comparing rates of extinction of a panel-

[8] Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, Stamford, Conn., American Cyanamid Co., 1949, 1-16.

pushing response in two groups of rats analogous to the present Groups *Cn* and *Rv*, found a significant difference opposite in direction to that observed in the present experiment.) Although the differences observed between Groups *Rv* and *Sa* did not attain significance, the order of the three groups with respect to rate of extinction was precisely that which was predicted by the discriminational hypothesis (*Cn—Rv-Sa*) and opposed to that predicted by reinforcement theory (*Rv—Sa-Cn*).

As previously noted, the analysis in terms of the discrimination-hypothesis assumes that rats learn not only about the occurrence of non-reinforcement, but also about specific stimuli associated with non-reinforcement. This interpretation is supported by the curve of Group *Sa* during extinction and also by the post-extinction curve of Group *Rv*. In both cases the experience of non-reinforcement in a goal-box previously associated with non-reinforcement led to a significant *decrease* in running time, while non-reinforcement in the positive goal-box (Group *Rv*), did not have such an effect. It appears, therefore, that non-reward accompanied by stimuli previously associated with non-reward may actually *strengthen* a response-tendency by virtue of its congruence with training conditions. The presentation during extinction of stimuli previously associated with reward does not strengthen the response-tendency when it is incongruent with training experiences. Both of these effects have been observed in previous experiments.[9]

The ultimate cause of the discrepancy between the present results and those reported earlier by Sheffield and Rubin remains unknown. The discrepancy may be due in part to the nature of the response studied, since it was not the same in the three experiments. The response-sequence required in the present experiment may perhaps be somewhat more reliably measured than running alone or panel-pushing. It should be noted, however, that Weinstock, in a recent runway experiment with rats, also failed to confirm Sheffield's results.[10] Using widely spaced practice, Weinstock found more rapid extinction of a running response after continuous as compared with partial reinforcement. Such results cast doubt on Rubin's data as well, since distribution of practice was supposedly one of the conditions essential to his unique results. Weinstock accounts for his findings in terms of the habituation of competing responses on non-reinforced trials in the partially reinforced groups. He assumes that non-reward elicits

---

[9] Bitterman, Feddersen, and Tyler, *op. cit.*, 456-464; Elam, Tyler, and Bitterman, *op. cit.*, 381-384.
[10] Solomon Weinstock, Resistance to extinction of a running response following partial reinforcement under widely spaced trials, *J. comp. physiol. Psychol.*, 47, 1954, 318-322.

responses which are incompatible with those elicited by reward. The partial reinforcement procedure results in the habituation of these incompatible responses to so low a level that they occur with reduced frequency during extinction and therefore produce less decrement in the response-class which was conditioned. Consistently rewarded Ss have no opportunity for habituation of this kind, and the higher frequency of occurrence of the competing responses during extinction in these Ss results in a greater decrement in the response-class originally conditioned. One of the most obvious (among the many) shortcomings of this explanation is its inability to account for the greater resistance to extinction which follows random partial reinforcement as compared with alternating partial reinforcement.[11] In the experiments by Longenecker, Krauskopf, and Bitterman with human Ss, and by Tyler, Wortz, and Bitterman with rats, equal frequencies of reinforcement and non-reinforcement should have provided equal opportunity for the habituation of competing responses, yet marked differences in resistance to extinction appeared as a function of the pattern of reinforcement. Weinstock, who considers his experiment to have "reopened" the problem of partial reinforcement, is apparently unaware of these data. At best, of course, the concept of habituation begs the central question.

SUMMARY

Hungry rats were trained under conditions of distributed practice to traverse a runway and jump from its end to a goal-box for food. Group Cn was rewarded on each trial in a goal-box of a given color. Groups Rv and Sa were trained with random partial (50%) reinforcement. For Group Sa, a goal-box of the same color was used both on reinforced and on non-reinforced trials. For Group Rv, reinforced and non-reinforced goal-boxes differed in color (black and white). During extinction, the goal-box color associated with reinforcement during training was used for all groups. The three groups thus permitted comparisons analogous to those previously reported by Sheffield and by Rubin. The results obtained were contrary to those obtained in the earlier experiments. The two partally reinforced groups (Rv and Sa) extinguished significantly more slowly than the continuously reinforced group (Cn). Although the two partially rewarded groups did not differ significantly during extinction, Group Rv nevertheless tended to extinguish somewhat more rapidly than Group Sa. The results support a prediction based upon the discrimination-hypothesis, but fail to support a prediction based upon the concepts of stimulus-generalization and secondary reinforcement.

[11] Longenecker, Krauskopf, and Bitterman, *op. cit.*, 580-587; Tyler, Wortz, and Bitterman, *op. cit.*, 57-65.

EFFECT OF CONTOURS ON BINOCULAR *CFF* OBTAINED WITH SYNCHRONOUS AND ALTERNATE FLASHES

By GARTH J. THOMAS, University of Illinois

It has been known for many years that synchronous flashes delivered to corresponding retinal areas of the two eyes yield a higher *CFF* than is obtained with uniocular regard or with binocular regard and alternate flashes.[1] Perrin has suggested we call this phenomenon the "Sherrington effect."[2] A study of the role of interocular phase of flash in binocular flicker-discrimination has recently been reported,[3] the data being analyzed in terms of Crozier's theory of the flicker-contour.[4] No consistent change was found in the slope and abscissa-value of the point of inflection of fitted normal integrals in *Log I* as between uniocular or binocular contours with in-phase flashes and out-of phase flashes; but the theoretical F_{max} required for best fit was consistently higher for binocular contours with flashes in phase than for either uniocular contours or binocular contours with flashes out of phase in the two eyes. In Crozier's terms, the results indicate an increase in the frequency of contribution of hypothetical excitable elements to the visual effect in the binocular in-phase condition, with no increase in the total number of elements contributing. Such an interpretation is plausible, it was argued, if the existence of cells (presumably in the visual cortex) is assumed, on which excitation from corresponding areas in the homonymous hemiretinas converge, and if it is also assumed that excitation from either of the homonymous hemiretinas can stimulate these common cells, but that confluent excitation from both

[1] C. S. Sherrington, On binocular flicker and the correlation of activity of 'corresponding' retinal points, *Brit. J. Psychol.*, 1, 1904, 26-59.
[2] F. H. Perrin, A study of binocular flicker, *J. Opt. Soc. Amer.*, 44, 1954, 60-69.
[3] G. J. Thomas, A comparison of uniocular and binocular critical flicker frequencies: Simultaneous and alternate flashes, this JOURNAL. 68, 1955, 37-53.
[4] W. J. Crozier and Ernst Wolf, Theory and measurement of visual mechanisms: IV. Critical intensities for visual flicker, monocular and binocular, *J. gen. Physiol.*, 24, 1941, 505-535.

retinas enhances their contribution to the final visual effect, *e.g.* increases their frequency of firing.

One may note parenthetically that Cajal postulated such hypothetical cells in the cortex, which he called "isodynamic" cells, and which receive excitation from corresponding areas in homonymous hemiretina.[3] Thus he interpreted the functional significance of the crossed and uncrossed projection of the retinas in the visual cortex of each hemisphere to explain binocular fusion. Sherrington argued that a mechanism like that postulated by Cajal could not form the basis of binocular fusion because phase of flash has such a small effect on binocular *CFF*. Instead, he concluded that processes from homonymous hemiretinas remain physiologically independent and that binocular unification takes place *psychically*. He suggested that the function of the partially decussated neuro-optic pathways with representation of parts of each retina in the same hemisphere is not to subsume binocular fusion but to facilitate coördination of conjugate and fusional eye-movements.[6] Sherrington's interpretation is supported by the fact that this partial decussation of pathways at the chiasma is found only in mammals, and that only mammals possess conjugate eye-movements.[7] Certainly when stimuli of identical quality, intensity, and timing fall on corresponding retinal areas, the visual impressions from either eye alone or both eyes together are exceedingly equivalent. Nonetheless there is some binocular interaction in respect of sensitivity to flicker as a function of the relative timing of the flashes delivered to the two eyes. The very speculative hypothesis was suggested in a recent paper that the slight increase in binocular *CFF* for synchronous flashes is not based simply on brightness-summation, but derives instead from interaction of contour-forming processes in the cortical cyclopean retina.[8] Fry and Bartley have shown that the processes subserving border-contrast influence the binocular interactions involved in Fechner's paradox.[9] Bouman has shown that simultaneous and successive incremental thresholds measured in one eye are completely unaffected by stimulation of the other eye as long as edge-contours of the stimuli do not fall on corresponding retinal points. If binocular fusion occurs, the component stimuli in the separate eyes become indistinguishable and it is impossible to measure thresholds in one eye only, *i.e.* the visual impressions from the two eyes are completely unified. Even when the stimulus in one eye overlaps completely but extends beyond the stimulus in the other eye, thresholds in one eye are unaffected by stimulation in the contralateral eye as long as the borders of the two stimuli are not on corresponding areas.[10]

The hypothesis suggested here is that the relatively small changes in flicker-sensitivity associated with phase of flash depend on the increased

[3] Santiago Ramon y Cajal, *Histologie du Système Nerveaux de l'Homme et des Vertébrés*, 2,1911, 880.
[6] Sherrington, *The Integrative Action of the Nervous System*, 1906, 354-386.
[7] G. L. Walls, *The Vertebrate Eye and Its Adaptive Radiation*, 1942, 319-331.
[8] Thomas, *op. cit.*, 51-52.
[9] G. A. Fry and S. H. Bartley, The brilliance of an object seen binocularly, *Amer. J .Ophthal.*, 16, 1933, 687-693.
[10] M. A. Bouman, On foveal and peripheral interaction in binocular vision, *Werkgroep Waarneming*, Rapport No. WW-1953-11.

output of Cajal's hypothetical isodynamic cells when they are stimulated by simultaneously confluent excitation from corresponding retinal areas, but that these hypothetical cells are primarily involved in the process subserving the formation of binocularly fused contours in the cortical cyclopean retina. If such is the case, increasing the amount of edge or contour that falls on corresponding retinal areas should enhance the Sherrington effect, as compared with that obtained when the contours fall on non-corresponding retinal areas. The experiment reported was carried out in order to test this hypothesis.

Apparatus. The apparatus has been described previously in some detail.[11] Briefly, two beams of light originating from a common source are interrupted in focal planes by sectored disks with equal open and closed sectors. The beams pass through separate sets of wedges, balancers, and filters for independent control of intensity, and fall on two screens of milk-glass. Apertures in thin metal screens placed in front of the milk-glass are viewed, one by each eye, through low-power microscopes. Identical grid-patterns, photographically produced on film, are mounted in each aperture. The lines of the grid are oriented vertically in the left field and in the right field; the pattern could be rotated 90° to orient the lines either horizontally or vertically. The stimulus-fields are circular in shape with diameters subtending a visual angle of 6.6°. The grid-lines, which subtend an angle of about 14' wide, are spaced 1.15° from center to center. Five lines are visible. Viewed binocularly, the stimulus-fields fuse and are seen as one spot of light crossed by thin lines in an otherwise dark field. When one grid is horizontal and the other vertical, the two stimulus-fields still fuse into a single spot of light, but marked binocular rivalry is engendered in respect of the grid lines. Small points of red light, one reflected into the center of each grid, serve as fixation-points. Care is taken to center *O*'s pupils with respect to the exit-pupils of the microscopes, and his head is maintained in position by means of wax bite and a forehead-rest.

Subjects. The two *O*s, *TS* and *GT*, acted alternately as *O* and *E*. Both have had considerable experience at making flicker-discriminations.

Procedure. The apparatus does not provide a control of luminance common to the two stimuli. The stimuli were therefore equated in luminance at steps of 0.5 log-unit (mL.) throughout the range of about 6 log-units by means of careful measurements with a Macbeth Illuminometer. At each level of flash-intensity, thresholds in flash-rate for detection of flicker were measured under the following conditions of viewing: Uniocular thresholds were determined for each eye. Two binocular thresholds were determined with synchronous flashes in the two eyes, once with both grid-patterns oriented vertically, and once with the left grid vertical and the right grid horizontal. Similarly, two binocular thresholds with alternate flashes were determined, one with each of the two arrangements of the grid-patterns. Thus, at each level of flash-luminance, thresholds were measured under six conditions of viewing. At each session *O* was dark-adapted, and all measurements for

[11] Thomas, The effect on critical flicker frequency of interocular differences in intensity and in phase relations of flashes of light, this JOURNAL, 67, 1954, 634.

TABLE I

UNIOCULAR CFF, STRIPES VERTICAL FOR BOTH EYES

Log I_{mL}	TS		GT	
	Right	Left	Right	Left
−3.12	—	—	5.9	6.0
−2.62	9.4	9.0	7.7	8.3
−2.12	9.9	10.1	9.4	9.4
−1.62	12.8	12.4	11.7	11.7
−1.12	13.6	13.0	13.0	13.5
−0.62	16.4	17.0	17.1	16.7
−0.12	21.9	21.5	21.9	21.8
+0.38	28.7	28.6	27.8	27.7
+0.88	33.4	33.6	33.2	32.9
+1.38	38.4	38.0	37.7	37.1
+1.88	44.3	43.8	44.6	44.9
+2.38	46.8	46.8	50.6	49.4
+2.88	50.9	51.1	53.3	53.3

a given level, 10 judgments for each condition, were completed before going on to another level. Thresholds were determined in terms of critical flash-rate for detection of flicker by a modified method of limits, *i.e.* starting well above *CFF*, the flash-rate was reduced by small steps, which were delivered on signal from *O*, until threshold was reached and *O* indicated he had detected flicker. The signals were given by pressing microswitches, one in each hand of *O*, which activated lights visible to *E* on the control panel.

Results. The uniocular flicker-contours for both *O*s plotted in flashes per sec. and log mL. (curves fitted by inspection), are shown in Table I and
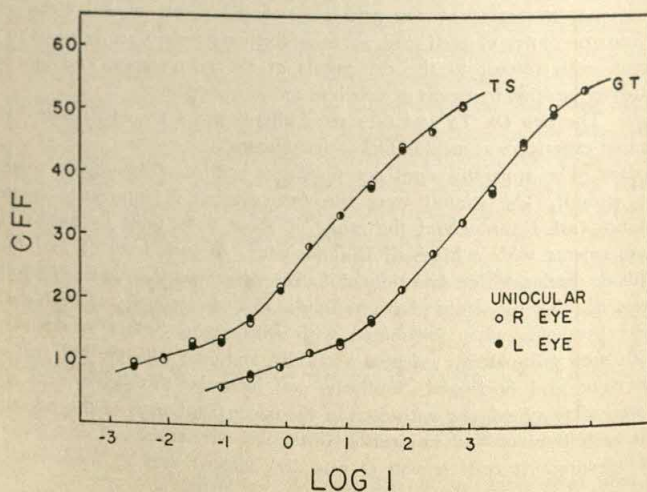
FIG. 1. UNIOCULAR FLICKER-CONTOURS FOR THE RIGHT AND LEFT EYES TAKEN SEPARATELY

The curve for *G T* has been shifted two log-units to the right.

TABLE II

Binocular *CFF*, Stripes Vertical for Both Eyes

Log I$_{mL}$	TS		GT	
	Flashes in phase	Flashes out of phase	Flashes in phase	Flashes out of phase
−3.12	—	—	7.0	6.1
−2.62	10.0	8.3	8.7	8.1
−2.12	11.9	9.6	10.3	9.1
−1.62	13.6	11.0	12.8	11.9
−1.12	15.0	12.2	13.6	13.2
−0.62	17.9	16.3	17.3	15.2
−0.12	23.3	20.9	23.9	22.1
+0.38	30.5	28.3	28.4	27.4
+0.88	35.3	32.6	34.7	32.1
+1.38	40.8	38.6	41.6	39.3
+1.88	45.2	42.3	45.5	44.0
+2.38	49.8	46.9	51.1	49.7
+2.88	53.3	50.6	54.8	53.0

Fig. 1. Abscissa-values are given for the curve of *TS;* the curve for *G T* has been shifted two log-units to the right. The grid-patterns in both right and left stimulus-fields were oriented vertically. Other data not presented here show, as would be expected, that orientation of the grid-lines has no effect on *CFF*-contours obtained uniocularly. There is no apparent systematic difference between right and left eye.

Binocular flicker-contours, obtained with alternate and synchronous flashes in the two eyes, are shown in Fig. 2 and Table II. Lines of the
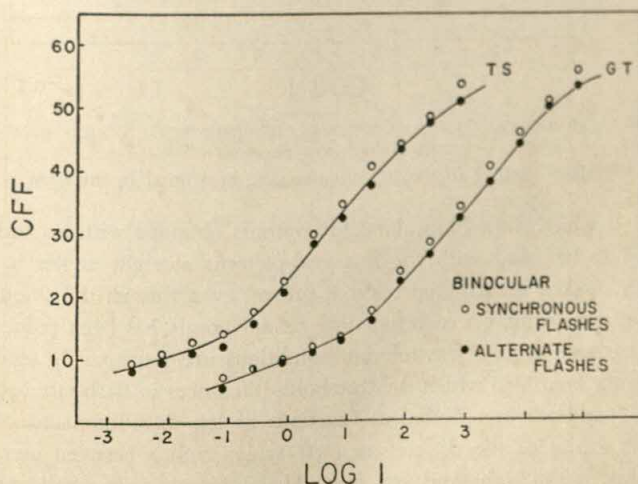
Fig. 2. Binocular Flicker-Contours Obtained with Synchronous and Alternate Flashes
Vertical grid-lines in both right and left stimuli.

grid-patterns were vertical in both stimuli. Both *Os* show a clear Sherring-
ton effect, in that *CFF* values obtained with synchronous flashes are con-
sistently higher than out-of-phase values. The curves drawn through the
data points were traced from those fitted to the uniocular data of Fig. 1.
They describe the out-of-phase data about as well as they do the uniocular
data. There is no consistent tendency over the whole contour for out-of-
phase flashes to result in lower critical flash-rates than those obtained with
uniocular stimulation.

Somewhat embarrasing to the hypothesis being tested are the curves
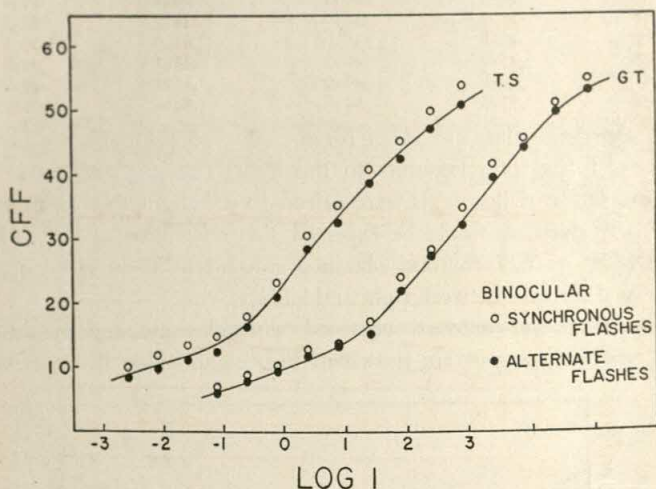
FIG. 3. BINOCULAR FLICKER-CONTOURS OBTAINED WITH SYNCHRONOUS
AND ALTERNATE FLASHES
Grid-lines vertical in the left stimulus and horizontal in the right.

of Fig. 3, which show binocular *CFF*-contours obtained with in- and out-
of-phase flashes, and with the two grid-patterns at right angles to each
other. A marked Sherrington effect is present even though the amount of
edge-contour falling on corresponding retinal points has been reduced.

Comparisons among the relevant conditions are more easily made by
considering Fig. 4, in which the threshold-differences in flash-rate between
various conditions are plotted as functions of log flash-luminance. Panel
A shows a plot of the differences (left *minus* right) between uniocular
thresholds of the right and left eyes. The differences are small and the
mean of the differences over the whole contour is not significantly differ-
ent from zero, which is to be expected because flash-luminance at each
level was carefully equated. Panel B shows the Sherrington effect obtained

TABLE III

Binocular CFF, Stripes Vertical for the Left Eye and Horizontal for the Right

	TS		GT	
Log I$_{mL}$	Flashes in phase	Flashes out of phase	Flashes in phase	Flashes out of phase
−3.12	—	—	7.2	5.4
−2.62	9.4	8.0	8.5	8.5
−2.12	11.0	9.6	9.9	9.4
−1.62	12.8	11.0	12.0	11.0
−1.12	14.2	12.0	13.6	13.1
−0.62	17.7	15.6	18.0	16.4
−0.12	22.3	20.7	23.9	22.4
+0.38	29.8	28.6	28.3	26.4
+0.88	34.6	32.5	34.0	32.3
+1.38	40.6	37.6	40.4	37.9
+1.88	44.2	43.8	45.8	44.0
+2.38	48.5	47.5	51.0	50.0
+2.88	53.6	50.8	56.0	53.3
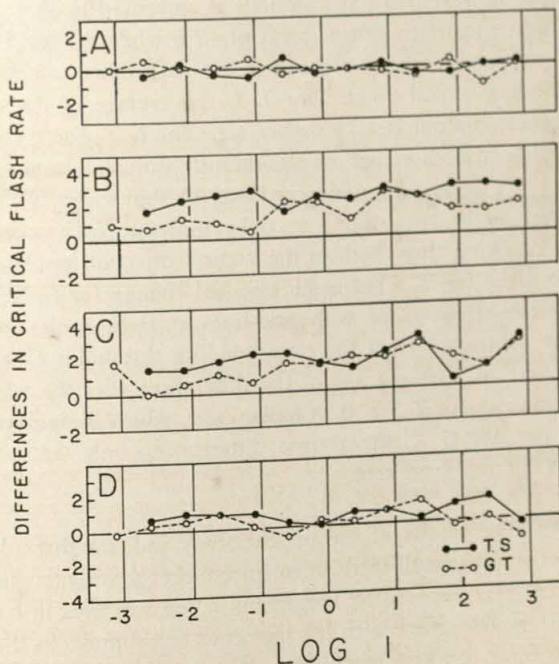
FIG. 4. DIFFERENCES IN CRITICAL FLICKER FREQUENCY

A, differences in *CFF* between right and left eyes; B, thresholds with synchronous flashes *minus* alternate flashes (grid-lines vertical); C, thresholds with synchronous flashes *minus* alternate flashes (grid-lines vertical in left stimulus and horizontal in the right stimulus); D, thresholds with synchronous flashes (grid-lines vertical) *minus* thresholds with synchronous flashes (grid-lines in the two stimuli perpendicular).

when the grid-patterns were both oriented vertically. At each level of flash-luminance the CFF obtained with alternate flashes was subtracted from the corresponding CFF obtained with synchronous flashes in the two eyes. Thus, deviations in the positive direction from the horizontal line originating at zero on the ordinate indicate the Sherrington effect. For *T S* (solid circles), the average increment in flicker-sensitivity along the whole contour is 2.42 flashes/sec.; for *G T* (open circles), the average increment is 1.44 flashes/sec.; and both values are significantly different from zero at less than the 1/10% level (*t*-test). These average values of the Sherrington effect are 1.69 times greater for *T S* and 1.26 times greater for *G T* than the average Sherrington effect obtained in an earlier experiment under comparable conditions but with homogeneous circular stimuli.[12] The results indicate that the presence of additional contours in the stimulus-patch tends to enhance the magnitude of the Sherrington effect. Panel C shows however, that synchronous as compared to alternate flashes cause very nearly as much enhancement of CFF when the two binocularly combined grid-patterns are at right angles to each other and thus are not on corresponding retinal points. For *T S*, the average of the differences along the whole contour is 1.75 flashes/sec.; for *G T*, the corresponding value is 1.42; and both values are significantly different from zero at less than the 1/10% level of confidence. Panel D shows clearly that setting the grid-lines in the two stimuli at right angles to each other as compared with orienting them both in the vertical direction results in less of Sherrington effect for *T S* but in no essential change for *G T*. Binocular in-phase thresholds obtained with grid-lines at right angles in the two stimuli were subtracted from the corresponding thresholds obtained with both grids vertical. The average of the differences over the whole range of flash luminance for *T S* is 0.58 flashes/sec., which is significant at the 1% level, but, for *G T*, the average difference is only 0.12 and is not significantly different from zero.

Discussion. The results of the present study indicate that addition of contours to the test-stimuli tends to enhance flicker-sensitivity when flashes are synchronous as opposed to alternating in the two eyes in comparison with the comparable effect with homogeneous test-stimuli. In the summation of binocular in-phase stimulation, it makes no difference, however, whether or not the added contours fall on corresponding retinal areas. The latter finding is embarrassing to the hypothesis, set forth earlier in this paper, that binocular summation occurs only when contour-processes in

[12] Thomas, *op. cit.*, this JOURNAL, 68, 1955, 37-53.

the cortical cyclopean retina are occasioned by corresponding contour-stimuli on homomymous hemiretinas because only then do the hypothetical isodynamic cells in the cortical retina (Cajal), which receive confluent excitation from both eyes, become involved in functioning as a final common path for binocularly fused contours. Even without the grid-lines, on the other hand, the stimulus-lights set up binocularly unified contour processes in the cortical cyclopean retina because of the contrast-borders of the test-patches themselves. Perhaps a more efficient evaluation of the hypothesis could be made by determining the effect on the Sherrington phenomenon of reducing rather than increasing the binocularly unified contours in the test-stimuli. Decreased binocular interaction (Sherrington effect) in respect of flicker should result under such circumstances. Equivalent stimulus-fields with no contours could be presented by providing the eyes with separate *Ganzfelder* which could be flickered synchronously or alternately. Under such conditions the hypothesis would lead one to expect little or no binocular interaction as indicated by the Sherrington effect.

Summary. From a hypothesis suggested by results of an earlier study of the effect of phase of flash on binocular *CFF*, it was predicted that addition of contours to the stimulus-patches which fall on corresponding retinal areas would enhance the Sherrington effect, *i.e.* the superiority of flicker-sensitivity when flashes are delivered synchronously to the two eyes as compared to *CFF* obtained with alternate flashes. An experiment is described in which binocular *CFF* with both in- and out-of-phase flashes in the two eyes was determined over a range of about 6 log-units of flash luminance under two conditions. In one condition the stimulus-patch viewed by the left eye and that viewed by the right eye contained grid-patterns of vertical lines arranged to fall on corresponding retinal areas when the stimuli were unified binocularly. In the other condition the grid-lines of the two stimulus-patches were perpendicular to each other. The stimulus-patches were circular in shape and identical in size and luminance. The results show an increase in the magnitude of the Sherrington effect when the stimulus-patches contain additional contours as compared with that reported in an earlier paper and obtained under comparable conditions. The arrangement of the lines of the grids (either vertical for both eyes, or vertical for the left and horizontal for the right) had, however, no effect on the magnitude of differences between *CFF* with in-phase flashes and *CFF* with out-of-phase flashes. An alternative test of the hypothesis is proposed.

ON THE DERIVATION OF EQUATIONS RELATING DATA OF FRACTIONATION AND CONSTANT-SUM TO PHYSICAL SCALES OF MEASUREMENT

By Katherine E. Baker and Frank J. Dudek, University of Nebraska

One aspect of the problem of psychological scaling concerns the derivation of equations for the relationships between psychological and physical dimensions. Several authors have dealt with this aspect of scaling and have developed different procedures for treating data of fractionation and constant-sum.[1] This paper considers a procedure for deriving equations which is applicable to both types of data. The problem of comparing scales obtained in different experiments by different methods makes clear the need for a logically sound procedure by means of which all scales can be put on a common basis. The present concern over problems of deriving equations arose in connection with a previous article showing considerable differences in seven psychological weight-scales obtained by different methods,[2] and the adequacy of the proposed solution is illustrated in terms of a comparison of various procedures as they apply to those seven sets of data. While the discussion takes up weight-scales specifically, the logic of the derivation of equations and the import of the comparison of procedures is more general.

Procedures. Fractionation provides a set of physical measurements of a standard weight, W, and the corresponding judged-half weight, W_h. The function relating W_h to W contains the fundamental facts about the way in which psychological weight is related to physical weight, but it provides no psychological scale of weight. The implied psychological scale must be derived from the function.

Although the W-W_h data only approximate a smooth curve, it seems reasonable to fit the data with the assumption that deviations are due to random experimental error. While plots of the raw data are usually curvilinear, log W_h plotted against log W approximates linearity fairly well, and thus data so expressed may be fitted by the method of least squares to get a straight-line equation in the form,

$$\text{Log } W_h = n \text{ log } W + \text{log } m \quad \dots\dots\dots\dots\dots [1]$$

where m and n are constants.

[1] R. S. Harper and S. S. Stevens, A psychological scale of weight and a formula for its derivation, this JOURNAL, 61, 1948, 343-351; J. C. Armington, A note concerning the *veg* scale of apparent weight, this JOURNAL, 66, 1953, 304-306; J. P. Guilford, *Psychometric Methods*, 1954, 208-221.
[2] K. E. Baker and F. J. Dudek, Weight scales from ratio judgments and comparisons of existent weight scales, *J. exp. Psychol.*, 50, 1955, 293-308.

From Equation [1] it is possible to state for any W what its W_λ should be. Thus, we may arbitrarily choose a unit and find the physical weights which correspond to multiples of this unit. Following the suggestion of Harper and Stevens, the psychological weight unit may be called a *veg* (V), which they define as the psychological weight corresponding to 100 gm. of physical weight.[3] Entering $W = 100$ gm. in Equation [1], we can solve for W_λ, which is the physical weight corresponding to $0.5V$, since V is defined as the psychological weight of 100 gm. The newly found W corresponding to $0.5V$ may be used in the equation to obtain W_λ which corresponds to $0.25V$, and so forth. It is possible to get Ws corresponding to V-values above 1.00 by entering $W_\lambda = 100$ in the equation and solving for W, or the weight corresponding to $2V$. Ws corresponding to 4, 8, or $16V$ may be obtained in similar fashion. These V-values, when plotted against corresponding W-values, show the relation between scaled psychological weight and physical weight, but it is not too satisfying to stop here, for an equation relating V to W would provide a much more general and exact description of the relationship.

Guilford has observed that log V plotted against log W seems to be a linear function and his formulation relating V to W consists in determining a least-squares fit for log data in the form,

$$\text{Log } V = a \log W + b \quad\quad\quad\quad\quad\quad [2]$$

where a and b are constants.[4] We shall see that Equation [2] is not applicable in a general enough way to stand as more than an approximation of the relationship. While this solution is the simplest in form, it appears that the assumption of a linear relation existing between log V and log W is not valid.

In 1953, Armington noted that the mathematical derivation by Harper and Stevens of their V-W equation involved operations not generally permissible.[5] As a substitute, he proposed that the solution be formulated in terms of progressions describing the *veg*-scale and the physical scale. Actually the procedure illustrated above for finding Vs and their corresponding Ws gives the required progressions, but Armington expressed the progressions in such general mathematical terms that they might be evaluated for any given term. Further, he was able to write an equation to describe exactly the relation between the *veg* and physical scales, as follows:

$$\log V = [\log 2/\log n][\log(\log W_0{}^{1-n} - \log m) - \log(\log W^{1-n} - \log m)] + \log V_0 \quad\quad\quad\quad [3]$$

where $V_0 =$ the number of psychological units corresponding to a given physical weight, W_0, and m and n are the constants in Equation [1].

While Equation [3] is certainly more cumbersome than Equation [2], the point to be emphasized is that Equation [3] gives an exact fit while Equation [2] does not. Under certain circumstances, considered below, differences between the two procedures are not great, but in other cases the implications of the linear fit deviate greatly from the progressions upon which the fitted data are based. The precision

[3] Harper and Stevens, *op. cit.*, 343-351.
[4] Guilford, *op. cit.*, 208-221.
[5] Armington, *op. cit.*, 304-306.

of Armington's solution would seem to warrant the greater complexity of the final expression. Actually it is not difficult to solve Equation [3] for V with $V_o = 1$ and $W_o = 100$.

Deriving psychological scales for data obtained by the constant-sum method involves a procedure somewhat different from that used for data from fractionation. The end-result of the constant-sum method is a set of judged ratios for the scaled stimuli, and the reference-stimulus or unit, usually the smallest stimulus, may be chosen arbitrarily. To relate this scale to already existing scales, a common unit must be used. In the present case 100 gm. $(= V)$ is the unit. If the scaled stimuli include a 100-gm. weight, judged relationships of other stimuli to this weight may be considered to provide V-values directly. For the case where no weight in the series corresponds exactly to this unit-weight, the original scale must be transformed. For example, if the smallest stimuli in a series are 40 and 150 gm., with scale-values of 1.00 and 3.2, respectively (where 40 gm. is the unit), a 100-gm. weight, by interpolation, has a scale-value of 1.2. To transform the original scale to one with 100 gm. as the unit, all original scale-values are divided by 1.2.

Again it is desirable to write the equation relating V to W, and there are three ways in which that may be done. First, as Guilford noted, V-values may be obtained from the original scale as described above, and a plot of log V against log W appears to be roughly linear. Thus, these V-data in log form may be fitted by a straight line. A second procedure rests upon the fact that a plot of obtained scale-values against W allows various Ws and their judged W_hs to be read from the curve, or found by arithmetical interpolation. With such a set of Ws and W_hs, Equation [1] may be applied, and (parallel with the treatment of the data of fractionation) specific V-W data may be calculated and fitted as in Equation [2]. The form of Equation [2] is the same as that of the equation obtained by the first procedure, but the constants will quite likely differ. The third procedure also makes use of the linear function relating log W to log W_h. With the constants m and n from Equation [1], Equation [3] may be solved to provide an exact description of the two progressions, as with fractionation-data.

Comparisons of alternative procedures. Any choice among the proposed equations for purposes of comparing scales can be justified by comparing the alternative procedures in terms of the fits provided for available data on the judgment of weight. Data from fractionation have been obtained by Harper and Stevens,[6] by Guilford,[7] and by Joy,[8] with a somewhat different experimental technique used in each case. Constant-sum data have been presented by Guilford,[9] and by Baker and Dudek.[10] In these studies, too, details of procedure differed, with ratios estimated directly by Baker and Dudek.

---

[6] Harper and Stevens, *op. cit.*, 343-351.

[7] Guilford, *op. cit.*, 208-221.

[8] Unpublished study; for further information see Baker and Dudek, *op. cit.*, or communicate with the authors.

[9] Guilford, *op. cit.*, 208-221.

[10] Baker and Dudek, *op. cit.*, 294-296.

Table I contains the equations derived by the alternative procedures for the seven experiments cited. Odd-numbered equations show the solutions of Equation [3] for V; even-numbered equations are solutions of Equation [2] for 10 datum-points (0.125, 0.25, 0.5, 1, 2, 4, 8, 16, 32, 64V) calculated by Equation [1]. For the constant-sum data, the equations desig-

TABLE I

EQUATIONS FOR Veg-WEIGHT RELATIONS DERIVED BY DIFFERENT PROCEDURES

Even numbers from Equation [2], prime numbers from original scales, and odd numbers from Equation [3].

Data	$\log V$	V
Fractionation (Harper & Stevens)	[4] $1.9428 \log W - 3.9192$	[5] $[0.1420/\log W^{0.0397} + 0.0626]^{-17.11}$
Fractionation (Guilford)	[6] $1.2346 \log W - 2.4692$	[7] $[0.2489/\log W^{-0.0105} + 0.2699]^{6^{e.31}}$
Fractionation (Joy)	[8] $1.0294 \log W - 2.0632$	[9] $[0.2895/\log W^{0.0051} + 0.2793]^{-135.r0}$
Constant sum 2 stimuli (Guilford)	[10] $1.0278 \log W - 1.9718$ [10'] $1.0637 \log W - 2.0982$	[11] $[0.3365/\log W^{-0.0850} + 0.5065]^{8.50}$
Constant sum (Dudek & Baker)	[12] $.9191 \log W - 1.6347$ [12'] $1.0707 \log W - 2.2022$	[13] $[0.4147/\log W^{-0.2007} + 0.8161]^{3.79}$
Direct estimation of ratios (Dudek & Baker)	[14] $.9759 \log W - 1.7038$ [14'] $1.2954 \log W - 2.6775$	[15] $[.3976/\log W^{-0.2457} + 0.8890]^{3.15}$
Constant sum 5 stimuli (Guilford)	[16] $1.1760 \log W - 2.3468$ [16'] $1.1545 \log W - 2.2797$	[17] $[0.2589/\log W^{-0.0057} + 0.2703]^{121.87}$

nated by even numbers with primes are the least-squares solutions for $\log V$ against $\log W$ when the V-values are obtained directly from the original scales.

In Fig. 1, 2, 3, and 4, leg 10 V is plotted as a function of $\log W$ to avoid negative logarithms (while preserving the shape of the curves). For the fractionation experiments (Fig. 1), the solid line is Equation [3] and the dashed line is Equation [2] for the V-values indicated by datum-points. For the constant-sum experiments (Fig. 2, 3, and 4) the curved solid line is Equation [3] again; the solid straight line is Equation [2] for the same 10 datum-points (not shown), and the dashed straight line is the fit for V-values obtained directly (open circles). To avoid superimposition in Fig. 2, Guilford's data have been moved 1 log-unit to the left.

The first feature of these plots to be noted is that Equation [3] provides a good fit of the data in all cases. The fit is exact for the calculated V-values, as it would have to be by the nature of the derivation, and V-values obtained from the scale-values determined by the constant-sum method are fitted closely by Equation [3].

In contrast, Equation [2] results in poorer fits. Fig. 2 and Fig. 3, es-
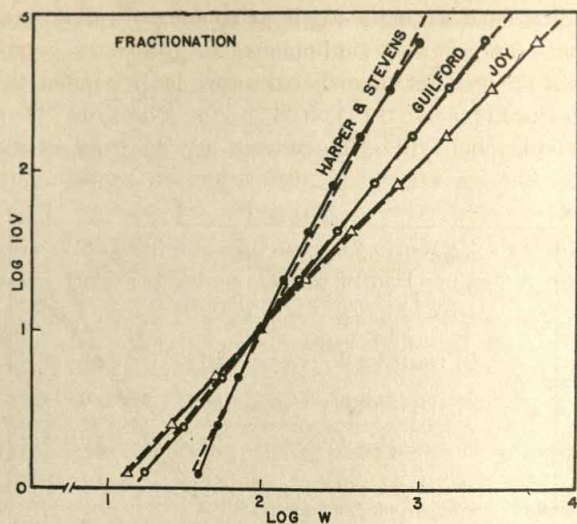
FIG. 1. LOG 10 V AS A FUNCTION OF LOG WEIGHT (IN GM.): FRACTIONATION
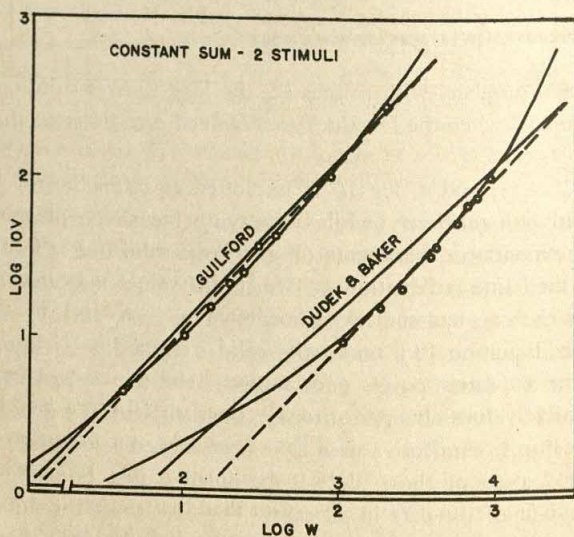
FIG. 2. LOG 10 V AS A FUNCTION OF LOG WEIGHT (IN GM.):
CONSTANT SUM, TWO STIMULI

pecially, indicate that the log V-log W relationship is curvilinear. As such, straight lines cannot fit and, furthermore, the greater the number of calculated V-values used to solve the equation, the less good the fit. For data of fractionation, the two procedures for fitting the obtained V-W relationship differ but little. This outcome may be traced to the fact that the slopes of the log W_h-log W relationships are nearly unity, and the
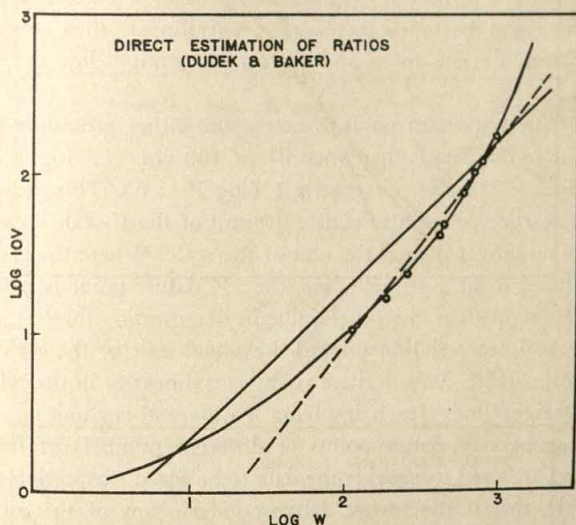
FIG. 3. Log 10 V AS A FUNCTION OF LOG WEIGHT (IN GM.): DIRECT ESTIMATION OF RATIOS

exponents in Equations [5], [7], and [9] are therefore quite large. The size of the exponent is determined by the $(\log 2)/(\log n)$-term in Equation [3] which approaches infinity as n approaches 1. Whether or not this state of affairs is typical of data of fractionation in general, or of data for weights only, has not been determined.

It is to be expected that a better fitting straight line will be obtained from a relatively restricted range of V-values than from a wider range if the relationship between V and W is curvilinear. Such appears to be the case for most constant-sum data where Equation [2] for a wider range (solid line in Fig. 2 and 3) fits less well than the straight line fitted to the original V-values covering a restricted range (dashed line in the same figures). It should be noted, however, that, even for the latter straight lines, the points which lie below the fitted line are predominantly in the middle of the range rather than randomly distributed, as would be ex-

pected if the trend were truly linear. When Equation [3] is available also, the fits of straight lines to the restricted range of original V-values appear less satisfactory, since extrapolations both above and below this range differ markedly for the two procedures. When an equation is given to describe a relationship, the usual interpretation is that the equation applies beyond the specific data; consequently the difference in the implications of the equations is important. As the exponent in Equation [3] becomes smaller, the trend becomes increasingly curvilinear, thus providing expressions which deviate more and more from straight-line fits of the V-data.

An especially important point to note about either procedure for fitting straight lines to the data is that when $W = 100$ gm. (*i.e.* log $W = 2$) the corresponding V-value is not exactly 1 (log $V = 0$). This means that in spite of the earlier decision to equate the unit of the V-scale with 100 gm., the relation no longer defines the unit of the scale. Where the least-squares method is used to fit a straight line, the $1V$ datum-point is neither more nor less important than any other value in determining the line. It is clear that since the linear relation implies the actual unit of the scale, the unit cannot remain stable. Where there is any curvilinearity in the relationship, different straight lines (each implying a different *veg*-unit) result when different numbers of datum-points or different numbers of steps in the progressions are used to provide the data to be fitted. Inspection will show, for example, that if the lowest datum-points in any of the constant-sum experiments were removed, the straight line fitting the remaining points would have a quite different slope. In contrast, Equation [3] always passes through the point $1V$, 100 gm.

Fig. 4, for the case where 100 points were divided among 5 stimuli, requires a few additional words of explanation. There is very little difference among the three curves shown, and all three appear to fit the data. The V-values depicted are not, however, those given in Guilford's report of the experiment. The point-assignment datum for 100 gm. was omitted here since there was reason to believe, as Guilford and Dingman also have noted,[11] that for this particular weight the data were especially divergent. This datum-point is a particularly important one in determining the constants of the equation which is derived from the original V-value data. If we accept the point-assignments for 40 gm. and 150 gm. as reliable, their judged relation to each other indicates a 1:3.68 ratio, which implies that a

[11] J. P. Guilford and H. F. Dingman, A validation study of ratio-judgment methods, this JOURNAL, 67, 1954, 395-410.

100-gm. weight should be perceived as 2.46 times as great as a 40-gm. weight. The point-assignments reported, however, lead to the inference that this ratio is only 2.06. The weight having this ratio to 40 gm. would be 83.5 gm., given the judged relation between 40 and 150 gm. Thus, to employ the reported data for the 100 gm.-weight to define the unit means that the V-values obtained are more like what the other data would indicate if V were to equal 83.5 gm. rather than the defined value of 100 gm.
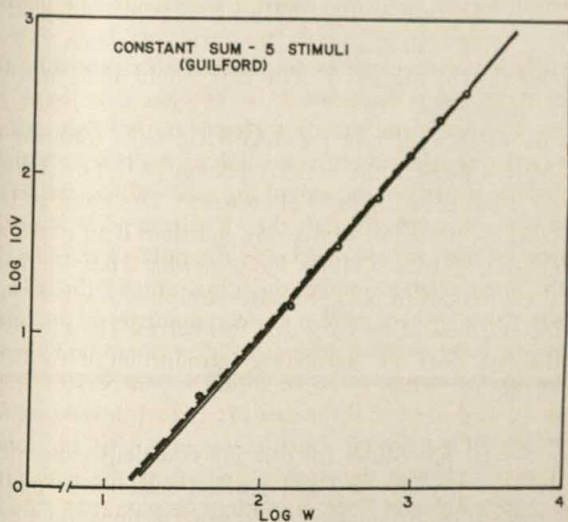
FIG. 4. LOG 10 V AS A FUNCTION OF LOG WEIGHT (IN GM.):
CONSTANT SUM, FIVE STIMULI

When the V-data that Guilford presents are fitted with a straight line, the value for V is out of line and the final curve fitted passes considerably above V-1 at 100 gm. If this one datum-point is omitted, the V-value for 40 gm. becomes 0.406 rather than 0.485, and the 2000-gm. weight corresponds to 34.64V instead of 41.63, as Guilford reports. If repeated measurements using the same experimental conditions that Guilford employed should show that his reported result for the 100 gm.-weight is reliable, then our values would need revision.

Discussion. With regard to the procedures employed to derive equations for weight-scales obtained by either scaling technique, Armington's solution appears the most adequate. It should be noted, however, that while

the information needed to write the required W-W_h equation for fractionation data does not involve additional mathematical treatments of the basic judgmental data, the same is not true of constant-sum data. In the latter case, one scale is already derived through several mathematical procedures before Armington's solution ever enters the picture. Baker and Dudek have shown that varying the procedure for deriving the original scale-values can affect the scale-values obtained by the constant-sum method.[12] It is also important to recognize that the procedure of halving scale-values involves the assumption that such half-values have a psychological meaning equivalent to judgments of one-half weight. That Equation [3] involving this assumption fits the original V-scale data would seem to attest to the validity of the assumption.

The data on weight-judgment presented yield quite different equations even when derived through the use of Equation [3]. The significance of these differences has been discussed elsewhere.[13] The point to be emphasized here is that the procedure of deriving equations is an important aspect of evaluating scaling data. Calling for somewhat complex and precise mathematical expressions of obtained relationships may appear pretentious in view of the exactness of the basic data themselves. The fact that error is present in the data, does not sanction the introduction of further error by imprecise treatments.

Summary. Several procedures for deriving equations expressing the relationship between psychological and physical scales were outlined. The procedure proposed by Armington for data of fractionation was found to be the best on logical grounds and on the basis of adequacy of fits for seven sets of weight-judgments. It was shown that data obtained by the constant-sum method could be treated in the same way with an added assumption, thus providing a means for achieving a common basis for comparing the two types of scaling-data.

---

[12] Baker and Dudek, *op. cit.,* 296-297.
[13] Baker and Dudek, *op. cit.,* 304-308.

SERIAL NON-RANDOMNESS IN AUDITORY DIFFERENTIAL-THRESHOLDS AS A FUNCTION OF INTERSTIMULUS INTERVAL

By WILLARD F. DAY, The Johns Hopkins University

Much interest has been shown recently in the apparent lack of independence of successive psychophysical judgments. When the serial order of the responses is analyzed, runs of similar responses are found which occur more frequently than would be expected in a random series.[1] This so-called *serial effect* has usually been studied under conditions which involve the repeated presentation of the same threshold-stimulus. An estimate is first made of the point at which 50% of the responses are correct, for the most part by a conventional method of constants or limits. Then a series of several hundred stimuli having this threshold-value is presented at a fixed rate, with an interstimulus interval typically on the order of several seconds. *O* is instructed to respond in some way whenever the stimulus is perceived, and the resulting response-series is analyzed for serial non-randomness in terms of the distribution of runs, dependent probabilities, serial correlation, or some other measure of sequential statistical organization.

At present the psychological basis of the effect is not precisely known. Under the usual experimental conditions, a number of factors probably contribute to the non-randomness of the response. Fatigue, boredom, wandering attention, physical discomfort, attitudes towards the experiment, and the like undoubtedly introduce a certain amount of organization into the series, but interest in the serial effect is chiefly centered about the possibility that it is caused by predictable changes in perceptual functioning. These changes may reflect orderly fluctuations in sensitivity characteristic of human performance at threshold, or they may reveal the presence of

[1] W. S. Verplanck, G. H. Collier, and J. W. Cotton, Nonindependence of successive responses in measurements of the visual threshold, *J. Exp. Psychol.*, 44, 1952, 273-282; Previous training as a determinant of response dependency at the threshold, *ibid.*, 46, 1953, 10-14; Collier, Probability of response and intertrial association as functions of monocular and binocular stimulation, *ibid.*, 47, 1954, 75-83; Collier, Intertrial association at the visual threshold as a function of intertrial interval, *ibid.*, 48, 1954, 330-334.

factors inherent in the relatively unique stimulus-conditions that make up the serial procedure.

In this experiment the serial non-randomness of threshold-series is investigated as a function of interstimulus interval (the interval of time separating successive stimuli within the series). If the stimuli are spaced at an interval greater than the average duration for which these serial factors act, the amount of non-randomness should be reduced. Such an average operation-time can be estimated by noting the shortest stimulus-separation at which the series appear to be approximately random. Interstimulus interval is varied in this study from 1.6 to 10.6 sec. under conditions that are generally associated with a large serial effect.

<center>EXPERIMENT</center>

Procedure. Especially convincing evidence of non-random threshold-performance has been obtained by Flynn for the auditory differential intensity-threshold.[2] The present experiment makes use of procedures similar to Flynn's. A continuous 1000 ∼ tone was presented to the right ear of each O at a sensation-level of 70 db. At a periodic rate the intensity of the tone was changed momentarily by the addition of a threshold increment, 0.1 sec. in duration. Following each increment the intensity returned to the 70-db. baseline. All stimuli within a particular session were given at the same interstimulus interval. O responded by pressing a key whenever he could detect a pulse in the loudness of the tone, and 'response' or 'no response' to each increment was recorded by the E.

At the initial session, each O was familiarized with the tone and increment until he said he had established a stable criterion for response. At the beginning of each subsequent session the 50% DL was determined by the method of limits with six ascending and descending runs. The intensity of the increment was fixed at this threshold-value, and a series of stimuli was presented at the appropriate interstimulus interval. To keep the length of all sessions somewhat comparable, the number of stimuli in the series was made a function of interstimulus interval. For intervals of 1.6, 2.1, 4.2, 7.1, and 10.6 sec. the series lengths were 600, 600, 450, 350, and 300 stimuli, respectively. The Os were discouraged from interrupting their observations in the middle of a series, but occasionally, during the longer sessions, 2-min. rests were given if requested.

Apparatus. The continuous tone and increment were generated by a Hewlett-Packard audio-oscillator (Model 200-B) and led through an amplifier to a single Permoflux headphone (Model PDR-10). The increment was produced without noticeable transients by changing the gain of the amplifier. The increment was introduced when a microswitch was closed by lugs spaced equidistantly on the periphery of a rotating aluminum disk. The disk was attached to the output-shaft of a constant-speed motor that revolved at a rate of one revolution per 21.1 sec.

---

[2] J. P. Flynn, Lack of randomness in sequences of auditory differential threshold data, *Amer. Psychologist*, 3, 1948, 254; J. P. Flynn, E. A. Jerome, and J. A. Moody, The intraserial dependence of psychophysical responses, *ibid.*, 7, 1952, 281-282.

The different interstimulus intervals were obtained by spacing 2, 3, 5, 10, or 13 lugs at equal separations on the disk.

Observers. The *Os* were five graduate students (men) in experimental psychology, and they were familiar in general with psychophysical procedures.

Design. The five conditions of interstimulus interval were arranged for the five *Os* within a Latin square. The experiment was carried out from Monday through Friday of one week, with one session a day scheduled for each *O*. The *Os* served each day at approximately the same hour.

RESULTS

Runs. The 25 response-series of this experiment were first tested for significant non-randomness by the statistics of runs.[3] The test compares the number of runs observed with the number of runs expected in a distribution of independent binary events. Where m represents the number of events in one class, n the number of events in the other class, and u the number of runs in the series, the mean and variance of the sampling distribution of runs for a sequence of independent binomial events having this m and n are given by the formulas $E(u) = (2mn/m + n) + 1$, and $\sigma_u^2 = 2\ mn\ (2mn - m - n)/(m + n)(m + n)(m + n - 1)$. If m and n are large, the distribution of runs is normal, and a test for the randomness of a particular series can be made by using z-scores in the conventional manner: $z = [u - E(u)]/\sigma_u$.

These relations show that serial z-scores do not provide a perfect measure of sequential non-randomness since they reflect only the probability with which a particular *number of runs* will occur in a series of independent events, *i.e.* the serial z is not a measure of all of the ways in which non-randomness can be present in a series. It is possible to construct highly organized series that have the same number of runs as that expected in a random serial arrangement. Nevertheless, chiefly because of the relative ease with which serial z-scores can be computed, the size of these scores is frequently used as a measure of the amount of serial effect. With sign disregarded, z-scores greater than 1.96 are significantly non-random at better than the 5% level. Negative scores indicate non-randomness in the direction of perseveration of response, while positive scores imply a tendency towards alternation.

Serial z-scores for the present data are given in Table I. These scores are based on response to the first 300 stimuli of each session. Only part of the data from the sessions having short interstimulus intervals is con-

---

[3] A. F. Mood, *Introduction to the Theory of Statistics*, 1950, 391-394. (A run is an unbroken sequence of similar responses. There are four runs in the series *RRRRNRRNNNNN*, *R* corresponding to 'response' and *N* to 'no response.')

sidered for analysis since serial z-scores are not comparable unless based on series of the same length. The amount of non-randomness seems clearly to depend on interstimulus interval when the data for all Os are pooled in medians and sums of ranks. The median z-scores for the shortest two intervals indicate on the average a much greater serial effect than for the longer intervals. Non-parametric analysis of the Latin square reveals a highly significant effect of interstimulus interval.[4] The scores for the shortest two intervals can be shown to be significantly larger than those for the longer intervals by the Mann-Whitney ranks test.[5]

A simple relationship between the amount of non-randomness and interstimulus interval does not, however, appear when the data for individual Os are considered. Although all of the series at the short intervals evidence a marked serial effect, the series for the longer intervals are random only for some of the Os. In the case of JH there is a highly sig-

TABLE I

SERIAL z-SCORES AND THEIR RANK-ORDER AS A FUNCTION OF INTERSTIMULUS INTERVAL

The number in parentheses after each z-score is the rank-order of that score relative to the other scores for the same O.

O	Interstimulus interval (sec.)				
	1.6	2.1	4.2	7.1	10.6
JH	− 8.980 (3)	−10.537 (1)	−8.697 (4)	−9.450 (2)	−5.072 (5)
CS	− 7.238 (1)	− 5.174 (2)	−0.697 (3)	+0.929 (5)	−0.115 (4)
RG	−11.516 (1)	− 9.462 (2)	−4.691 (3)	−0.374 (5)	−1.147 (4)
WG	− 7.310 (2)	− 7.827 (1)	+4.246 (5)	+2.466 (4)	+0.019 (3)
JV	− 8.100 (1)	− 6.527 (3)	−1.003 (4)	−6.546 (2)	+0.012 (5)
Median	− 8.100	− 7.827	−1.003	−0.374	−0.115
Sum of ranks	8	9	19	18	21

nificant serial effect at all intervals. For JV, even though the series at the 4.2-sec. interval is random, the serial effect for the 7.1-sec. interval is as large as it is for the 2.1-sec. interval. The data for WG are particularly interesting since two instances of significantly *too many* runs occur, indicating a tendency for response-alternation.

Serial correlation. Correlational methods also are often used to describe the statistical structure of threshold response-series. A coefficient of correlation can be a measure of serial non-randomness when the events paired

[4] M. Friedman, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *J. Amer. statist. Assoc.*, 32, 1937, 675-701.
[5] H. B. Mann and D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, *Ann. Math. Statist.*, 18, 1947, 50-60.

for correlation are members of the same series removed from each other by a constant lag. Since the series in this experiment are made up of discrete events of two categories it is appropriate to use the *phi*-coefficient of correlation,[6] a product-moment *r* for binary data.

Correlation functions for the 25 response-series were obtained by plotting *phi* as a function of lags from 1 to 20. Summary correlation functions for each interstimulus interval are given in Fig. 1.[7] They represent *phi* averaged at each lag for the five *O*s. Non-randomness within the series is reflected by the extent to which the functions deviate from the solid horizontal line at $\phi = 0$. Values of *phi* outside the area bounded by the two dashed horizontal lines show a significant correlation beyond the 1% level of confidence.

The functions for the two shortest intervals are plotted in the upper part of the figure. Those in the lower part are for the longer intervals. These average functions reveal a difference in serial effect for the two groups of interstimulus interval similar to that indicated by the serial *z*-scores. For the 1.6- and 2.1-sec. intervals the values of *phi* are highly significant through the first three lags. In contrast, the functions for the three longer intervals lie almost entirely within the range of non-significant correlation.

The dependence of the amount of non-randomness on the time between successive stimuli can be determined from Fig. 2. Here *phi* is plotted as a function of the temporal separation of the responses paired in correlation rather than as a function of lag. The points are *phi*s for Lag 1 and partial *phi*s for Lag 2 averaged for all *O*s. Comparable average *phi*-scores for lags greater than 2 show no significant correlation.

It is important to recognize clearly what these correlations mean. The values plotted for Lag 2 in this figure (separations of 2.1 and 4.2 sec.) are partial-correlation coefficients with the effects of correlation at Lag 1 held constant.[8] In this way the *phi*s are not inflated by significant correlation at Lag 1. In considering the interpretation of correlation functions it is necessary to keep in mind that the correlation at any lag may reflect to a certain extent significant correlation at shorter lags. The effects of such

---

[6] E. B. Newman, Computational methods useful in analyzing series of binary data, this JOURNAL, 64, 1951, 252-262.

[7] The power spectra for these average correlation functions have been analyzed in detail by R. P. Abelson (Spectral analysis and the study of individual differences in the performance of routine, repetitive tasks, Technical Report, Project Designation NR 150-088, Contract N6onr-27020, Princeton University, March 1953).

[8] The need for partialing out the effects of correlation at intervening lags was first pointed out to me by W. R. Garner.
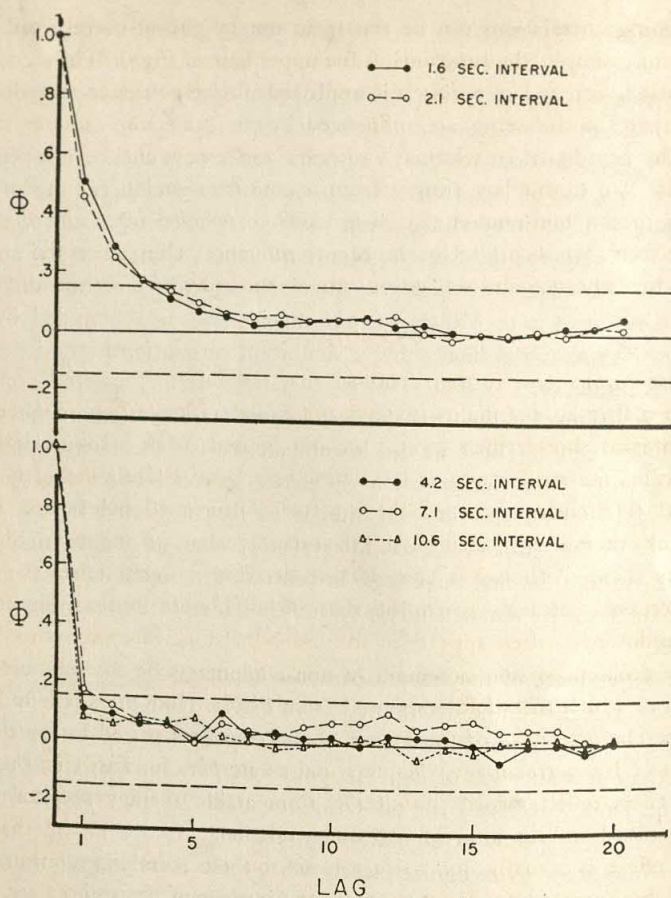
FIG. 1. AVERAGE *Phi*-SCORES AS A FUNCTION OF AMOUNT OF TIME BETWEEN
STIMULI PAIRED IN THE CORRELATION
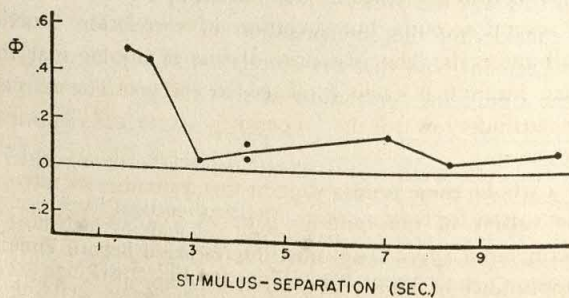
FIG. 2. AVERAGE SERIAL CORRELATION FUNCTIONS FOR FIVE DIFFERENT
INTERSTIMULUS INTERVALS

intervening correlations can be separated out by partial correlation. Consider, for example, the functions in the upper half of Fig. 1. The significant correlation seen at Lag 3 does not imply that the responses to stimuli three steps apart in the series are influenced by the same non-random factor, since the significant correlations at Lags 1 and 2 have not been taken into account. We cannot say simply from a consideration of Fig. 1 for how long a time a non-random factor is likely to operate, but from Fig. 2 it can be seen that such a factor appears to influence behavior on the average only where the stimulus-separations are on the order of 2 sec. or less.

DISCUSSION

These results give further evidence that response to the serial presentation of a threshold-stimulus is likely not to be random when an interstimulus interval shorter than several seconds is used. With a longer interval, significant non-randomness is less likely to appear. Postulation of a more general relationship between the amount of non-randomness and interstimulus interval is probably unwarranted, at least on the basis of data such as these. Although a considerable decrease in serial effect is found for increased intervals when the data of all Os are pooled, significant non-randomness often appears in the individual long-interval series.

The problem is not, however, so much to determine interstimulus intervals at which threshold response is likely to be random as it is to identify the characteristics of the human that account for the statistical organization of his performance. Of particular interest is the possibility that the serial effect reflects orderly changes in some aspect of the perceptual process involved in this kind of threshold-judgment. To the extent that the serial effect is actually due to factors of this type, the size of the effect would be expected to diminish with an increase in the interval of time separating correlated responses. Fig. 2 shows that just such a reduction in the amount of correlation occurs when the interstimulus interval becomes larger than several seconds, but a consistent behavioral picture does not appear at all intervals. The instances of non-randomness at the longer intervals may mean that factors related to the motivation, preliminary training, or attitude toward the experiment were not sufficiently controlled in this study.

Taken as a whole, these results suggest that a number of factors actually underlie the variety of non-random patterns that are loosely subsumed under the term *serial effect*. The most interesting aspect of the data is the consistent appearance of highly structured behavior at the short interstimulus intervals. Instead of hypostatizing too rigidly the concept of serial

effect at this point, it seems reasonable first to attempt to describe more precisely the range of conditions under which such consistent statistical control is apt to appear in threshold-behavior.

SUMMARY

Several recent experiments have shown that response to long series of threshold stimuli is not likely to be random in a sequential sense. Similar responses occur more frequently in long runs than would be expected from a distribution of serially independent events. The present experiment was to determine whether this effect could be reduced by increasing the length of time between successive members of the stimulus-series.

A continuous 1000∼ tone at a comfortable intensity was presented monaurally to Os who were instructed to respond by pressing a key whenever they could detect an increment in the loudness of the tone. From 300 to 600 increments in intensity, each 0.1-sec. in duration and of differential-threshold magnitude, were regularly added to the continuous tone at a fixed interstimulus interval. A response of 'yes' or 'no' was recorded for each increment. Five conditions of interstimulus interval, 1.6, 2.1, 4.2, 7.1, and 10.6 sec., were varied for five Os in a Latin-square design. The responses were analyzed for non-randomness by the runs-test and functions of serial correlation. The results show on the average a decline in the amount of non-randomness for interstimulus intervals larger than several seconds, but the data of individual Os often manifest marked statistical organization even at the long interstimulus intervals.

CHANGES IN THE MEMORY-TRACE FOR PERCEIVED FORMS WITH SUCCESSIVE REPRODUCTIONS

By EDWARD L. WALKER, University of Michigan,
and JOSEPH VEROFF, Princeton University

The principle that there is dynamic activity within the individual memory-trace has been espoused by Wulf and Koffka among others.[1] Many sub-principles have been offered to describe the nature of the systematic changes that occur. A short and simple list includes: (a) *normalization* or *closure,* which consists of changes that approach a more simple or a more familiar figure; (b) *pointing* or *sharpening,* which consists of exaggerations of certain characteristics of the figure; and (c) autonomous changes within the memory-trace as a result of intrinsic stresses of the visual pattern. It is not at all clear, however, when and under what conditions each should occur.

Typical results using the method of reproduction are those by Allport which show that children tend to make asymmetrical figures more symmetrical (normalizing) with repeated reproductions, and those by Gibson which show that adults tend to complete originally incomplete figures (closure).[2]

In reviewing the evidence for the Gestalt position, both Woodworth, and Hebb and Foord have pointed out defects in the method of reproduction.[3] Repeated reproductions induce new changes that interfere with the original trace so that the progress of the individual trace over time cannot be isolated. Hanawalt has found changes in recall by reproduction exaggerated as compared to recall by recognition.[4] Furthermore, he found that some errors in directly copying figures are the same as those that have been attributed to the effects of dynamic activity in the memory-trace.

---

[1] Friedrich Wulf, Ueber die Veränderung von Vorstellunge (Gedächtnis and Gestalt), *Psychol. Forsch.,* 1, 1922, 333-373. Quoted in K. Koffka, *Principles of Gestalt Psychology,* 1935, 493-506.
[2] G. W. Allport, Change and decay in the visual memory image, *Brit. J. Psychol.,* 21, 1930, 133-148; J. J. Gibson, The reproduction of visually perceived forms, *J. exp. Psychol.,* 12, 1929, 1-39.
[3] R. S. Woodworth, *Experimental Psychology,* 1938, 77-91. D. O. Hebb and E. N. Foord, Errors of visual recognition and the nature of the trace, *J. exp. Psychol.,* 35, 1945, 335-348.
[4] N. G. Hanawalt, Memory trace for figures in recall and recognition, *Arch. Psychol.,* 31, 1937 (No. 216), 5-89.

A fundamental assumption in the psychological theory that Hebb has proposed is that "the memory trace, the basis of learning, is in some way structural and static."[5] Hebb and Foord, using the method of recognition, obtained data on memory for patterns which were inconsistent with the Gestalt position.[6] In light of these results and the criticisms cited above, Hebb concluded that the concept of spontaneous changes in memory-traces should be abandoned, and that his fundamental assumption has not been challenged by any more recent research on the memory for form.

METHOD

While there seems little hope of using the method of reproduction and still escaping the criticism that the act of reproduction will influence the memory-trace, we thought we might be able to measure the changes introduced in the process of reproducing the figure, changes that might occur simply as a function of successive reproductions, and finally changes that occur as a function of the passage of time, in this instance a period of two weeks. If we were successful, we could then compare the magnitude and direction of such changes.

Stimulus-figures. It was hoped that, by careful choice of figures, we might gain some information concerning the conditions under which the principles indicated above might be expected to operate. The figures used in this study were two circles, two angles, and two quadrilaterals. One circle had a gap of 20° and the other 80°. Either could be opened or closed. The two angles presented were 55° and 150°. These also could be opened or closed. The two quadrilaterals were drawn to present two degrees of asymmetry and could be made either more or less symmetrical. Changes in the memory for the circle gap and the angle could be measured in degrees. For the changes in the memory for the quadrilaterals, we developed a distortion-ratio. This ratio was calculated by drawing a line from the upper left to the lower right corners of the quadrilateral (the shorter of the two alternatives), and then dividing the area of the larger resulting triangle by the area of the smaller.

Procedure. The two sets of figures used in this study may be seen in Fig. 1. A complete experiment was performed with Form A and another complete experiment was performed with Form B.

Each experiment was performed under two conditions, labeled 1 and 2. In Condition 1 the Ss were handed booklets in which there were six pages. There were figures printed on pages 1, 3, and 5, while the other pages were blank. The pages of the booklet were of double thickness that the Ss could not see the next figure in advance or the previous figure after the page was turned. They were informed that this was an experiment in the perception of form and immediately instructed as to procedure. At a 'ready' signal, they were given 5 sec. to look at the figure, were then instructed to turn the page, and then were given 15 sec. to reproduce it. They proceeded immediately to the second and then the third figure, with the whole process requiring 1 min.

[5] Hebb, *Organization of Behavior,* 1949, 12-13.
[6] Hebb and Foord, *op. cit.,* 335-348.

The booklets were hurriedly removed, and three additional blank pages were handed out. This process required 1 min. They were then given 15 sec. to reproduce a second time each of the three figures in order. Two weeks later they were again handed three blank pages and asked to reproduce the figures a third time. One group of Ss was given Form A and the other Form B.

Condition 2 was identical with Condition 1 except that two weeks separated the first and second reproductions, and the third reproduction followed the second by 1 min. A third group of Ss was given Form A in Condition 2 and a fourth group Form B in Condition 2.

Subjects. The Ss for this experiment were drawn from sections of the elementary psychology classes of the University of Michigan. A total of 108 Ss participated,

TABLE I

NUMBER OF Ss IN EACH CONDITION

Condition and Form	Total N	Stimulus-figure		
		Circle	Angle	Quadrilateral
1, A	23	23	20	23
1, B	16	15	13	12
2, A	29	29	27	24
2, B	26	26	21	19

but 14 were absent from the second session and were thus discarded. Further attrition occurred when some of the Ss were unable to recall one or more of the figures after two weeks. Only one S was unable to reproduce the circle with the gap, 13 were unable to reproduce the angle and 16 could not recall the quadrilateral. The number of Ss which remained in each group and for each figure may be seen in Table I.

RESULTS

The results may be seen in Fig. 2, 3, and 4, where the change which occurred in each figure on the first production by S (a comparison of the figure produced by S with the objective figure presented to him) is indicated by the broken line drawn from the horizontal line to a point indicating the mean value for the first production. This change will be called the productive error. Differences between the first two points in Condition 1 and the last two in Condition 2 are measures of the effects of an immediate reproduction, the reproductive error, and compare two figures drawn by each S. The two long lines in each case compare the figures drawn by the Ss two weeks apart, changes as a function of memory.

Productive error. In Fig. 1 it is clear that the Ss tended to widen the gap in the smaller circle and narrow the gap in the larger circle. The statistical significance of these changes was estimated by tabulating the number of Ss who opened the gap and the number who closed it and

running a chi squared against an expected even division of Ss. The resulting value of chi-squared for the circle with the smaller gap was significant at the 1/10% level, while that for the circle with the larger gap was significant only between the 20% and 30% levels.
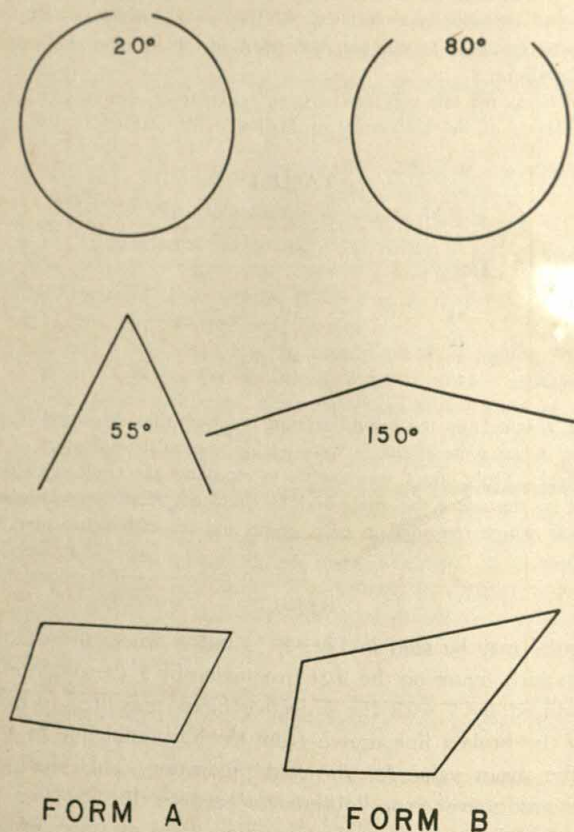
FIG. 1. FIGURES USED IN STUDY

The gaps in the circles are 20° and 80° and the circles were 60 mm. in diameter in the booklets. The angles are 55° and 150° with arms which were 55 mm. long. The quadrilaterals were 50 mm. from upper left to lower right corners.

In Fig. 2 it is clear there was a tendency to close the larger angle on the first reproduction, a change that was significant at the 1/10% level, but no tendency to change the smaller angle.

In Fig. 3 it will be seen that the tendency in both cases was to increase

the asymmetry of the figure. The significance levels were 1/10% for Form A and 10% for Form B.

Reproductive error. Examination of the figures indicates that the changes which occurred between two immediate reproductions by the *S* are essentially random and insignificant. The significance of these changes was computed by comparing the two immediately successive reproductions drawn by each *S* and tabulating the number of *S*s who distorted the fig-
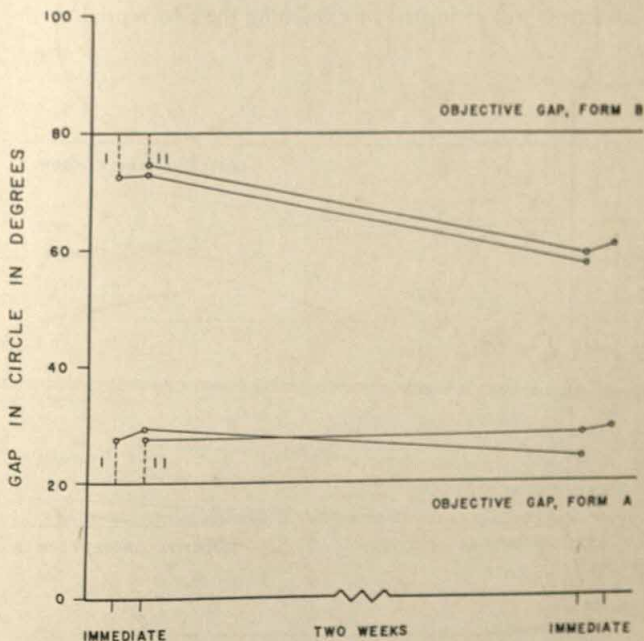
FIG. 2. DIAGRAM SHOWING PRODUCTIVE ERROR, REPRODUCTIVE ERROR AND CHANGES AS A FUNCTION OF TIME IN DRAWINGS OF TWO CIRCLES WITH GAPS

ures. Chi squared was run on these tabulations comparing them to an expected even division of *S*s. There is only one apparently significant change in Fig. 3 where there is a fairly large increase in distortion between two successive reproductions for one group of *S*s. This is the only difference to reach the 5% level or beyond, and it was significant beyond the 1% level. There is no clear basis for the expectation of such a change, and it must therefore be assumed that it occurred by chance. This conclusion is supported by the fact that this was the condition (Condition 1, Form B

for the quadrilateral) in which there was the smallest number of Ss with usable data *i.e.* 12.

Changes as a function of memory. To infer that the difference between two reproductions by the same S separated by a period of two weeks is a function of changes in the memory-trace over that period, it is necessary to compare such changes with both productive and reproductive errors.

In the instance of the circles in Fig. 1, there seems a clear tendency to close the 80° gap in the circle with the passage of time. The significance of the differences was estimated by examining the two reproductions done
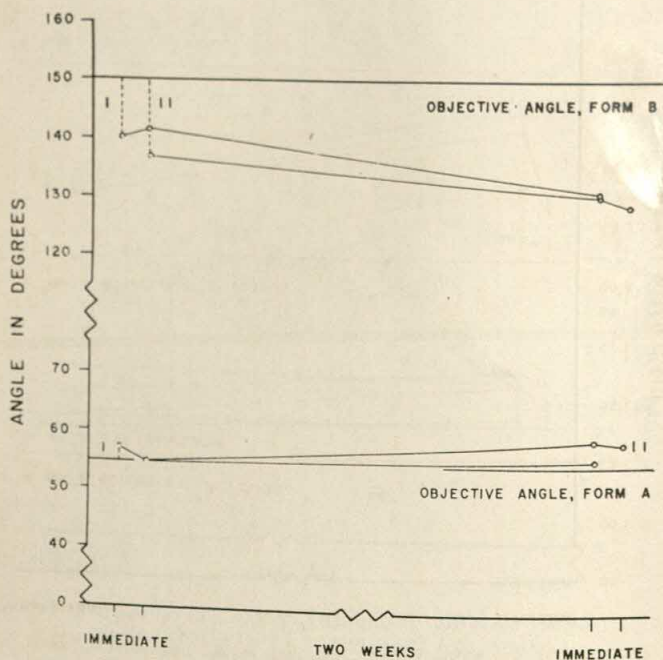
FIG. 3. DIAGRAM SHOWING PRODUCTIVE ERROR, REPRODUCTIVE ERROR AND CHANGES AS A FUNCTION OF TIME IN DRAWINGS OF TWO ANGLES

by each S and by tabulating the number of Ss who opened the gap as compared to the number who closed it, and running a chi squared of this division against an expected even division of Ss. The difference in one case is significant at the 2% level, at the 3% level in the other, and at the 1% level when the two groups are combined. The change was in the same direction as, and nearly twice the magnitude of the productive error,

and seems not to be attributable to successive reproduction alone, since the later factor produced no significant change. With the smaller gap (Form A), however, there was no significant change with the passage of time.

The changes occurring in the two angles (Fig. 2) with the passage of time present a very similar picture. The Ss tended to close the larger angle, a difference which is significant at the 2% level when both groups are combined. In this instance, however, the magnitude of the change is similar to that of the productive error. With the smaller angle (Form A),
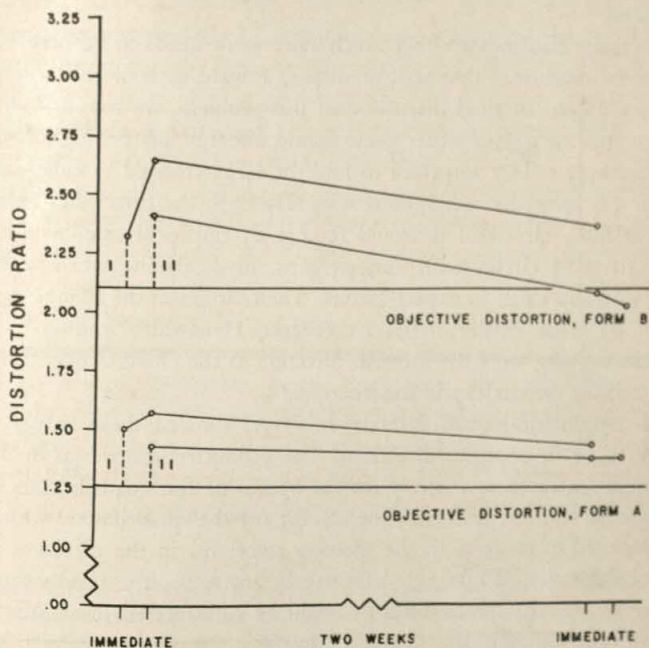
FIG. 4. DIAGRAM SHOWING PRODUCTIVE ERROR, REPRODUCTIVE ERROR AND
CHANGES AS A FUNCTION OF TIME IN DRAWINGS OF TWO
ASYMMETRICAL QUADRILATERALS

there was a slight tendency to open the angle which was significant at the 5% level with the two groups combined.

With the quadrilateral there was a clear tendency to make the more asymmetrical figure more symmetrical with the passage of time, a difference significant at the 1% level. With the less distorted figure, the tendency was in the same direction (toward more symmetry) but significant at only

the 20% level. With both, the direction of change with the passage of time was opposite to that of the productive error.

DISCUSSION

These results suggest that it is possible to use the reproductive method profitably, and obtain semi-independent estimates of productive error, reproductive error, and changes that occur as a function of the passage of time. The observed changes with the passage of time suggest that the conception of dynamic activity in the memory-trace cannot be completely abandoned.

The results also indicate that much more work needs to be done on the problem of conditions that will predispose toward each of the antithetical processes. In our original discussion of this problem, we found it possible to argue that the gap in either circle should undergo sharpening or closure. We were only mildly surprised to find the larger tended to close and the smaller was somewhat ambiguous with respect to the direction of change. It seems fairly clear that it would require an empirical examination of a wide variety of circles with various gaps, to determine when to expect sharpening and when to expect closure. The changes in the triangle seemed to parallel those in the circle. In general, Hanawalt's finding that the errors in copying were the same in direction as the changes over time was confirmed for the circles and the triangles.[7]

The asymmetrical quadrilaterals, however, showed these two types of change to be in opposite directions. The productive error was in the direction of increased asymmetry for the figures of the quadrilaterals. With the passage of time, however, the change for the quadrilateral which can be attributed to changes in the memory trace was in the direction of increased symmetry. While not definitive in any sense, the results for these figures suggest that great difficulty might be encountered in an attempt to specify independently the degree of intrinsic stress in a figure in such a manner as to predict the direction and degree of change over time.

---

[7] Hanawalt, *op. cit.*, 5-89.

TEMPERATURE CHANGES OF THE SKIN:
A FUNCTION OF INITIAL LEVEL

By ROBERT PLUTCHIK, Hofstra College, and
MILTON GREENBLATT, Harvard Medical School

Earlier clinical reports dealing with the relation between skin-temperature and emotion have reported that subjects (Ss) who reveal conflicting ideas tend to show decreases in their skin-temperatures (finger) and that those experiencing feelings of acceptance and reassurance by the therapist tend to show increases.[1] These results must be qualified, however, because the skin-temperatures of most of the Ss was 34°–35° C. at the start of the interviews, which is close to the upper limit of the normal range. A fall in temperature was, therefore, more likely than a rise.

If progress is to be made in our understanding of the relation between skin-temperature and emotion, the factors determining the physiological response must be isolated and identified. Some of these factors depend upon the stimulus; some upon the homeostatic, regulatory mechanisms in the body; and some upon S's perception of the situation.

If skin-temperatures vary over a relatively small range, then it is likely that a given stimulus applied to the body when the temperature is low will produce a quite different effect from that which would occur if the temperature were high.

The problem of this study was to discover how the temperature of the skin (finger) varies when the change occurs at different initial temperatures.

Procedure. Nine patients at the Boston Psychopathic Hospital, of various diagnostic categories, were used as Ss. Five psychiatrists working with them were requested to hold one or more interviews with these patients in the experimental laboratory. The laboratory room consisted of a bed in a sound-proofed, air-conditioned room, equipped with a pickup microphone, a one-way vision screen, and various electrical pickup-devices attached to S, including one for recording skin-temperature. The room-temperature was kept at approximately 27°C. (80°F.) and at a humidity of 65%. Previous work has shown that skin-temperatures are unaffected by humidity changes between 40% and 90%.[2]

---

[1] Robert Plutchik, The psychophysiology of skin-temperature: A critical review, *J. gen. Psychol.* (in press).
[2] R. Flecker, Skin-temperature as a psychophysical variable, *Aust. J. Psychol.*, 3, 1951, 109-116.

Most of the patients reclined on the bed during the interview, while a standard thermister taped to the forefinger of the hand picked up the temperature-changes. The continuous record obtained during the psychiatrist's interview of 1 hr. was condensed and plotted on graph paper and various values were then computed from it. Temperature changes of 0.5°C. or greater were alone noted and recorded in the analysis of the data. Changes of less than 0.5°C. were ignored.[3]

It was also necessary to adopt criteria of the end of a temperature change. Three were selected. A temperature change was considered terminated: (a) if there was no further change in temperature for 30 sec.; (b) if there was a change in direction of temperature of at least 0.1° over a half a minute; and (c) if the rate of change had slowed to 0.1 per min.

The magnitude, duration, and rate of any change meeting these criteria, and the initial temperature existing at the beginning of such a change, were also recorded. In all a total of 29 individual interviews were held, recorded, and analyzed. Two Ss *BW*, a 17-yr.-old girl, and *FL*, a 30-yr.-old man, provided 17 of these records. The analysis of the results is based chiefly upon the records of these two Ss.

Results. Two systematic errors were involved in the analysis of the results: (1) Since every S had different numbers of interviews, each contributed different amounts of information to the analysis; and (2) the records for the different Ss were not over comparable ranges of starting temperature. Although internal temperatures may remain nearly constant at 37°C., the temperature of the finger has been found to vary from 20° to 36°.[4]

Because of these errors, an attempt was made to assess the similarity of the data contributed by various Ss. The data of *BW* and *FL* were plotted according to the frequency of their starting temperatures and the magnitude, rate, and duration of the ensuing changes were recorded. Rises were treated separately from falls. Since their results were found to be remarkably similar—both showed almost identical frequency-distributions for the portion of the total range over which they coincided—their data were combined into a single distribution for each function.

For example, in Fig. 1, the distributions of *BW* and *FL* are compared for frequency of temperature-increase from given starting temperatures. These two distributions coincide with regard to maximal point and general shape. They differ only in the lower ranges reached. Fig. 2 shows their combined distributions for both rises and falls of temperature.

From these and the other combined distributions of the other Ss the following facts are indicated:

---

[3] Roderick Manzies, Conditioned vasomotor responses in human subjects, *J. Psychol.*, 4, 1937, 75-120; J. M. Steele, Fever in heart failure: Relations between the temperatures of the interior and the surface of the body, *J. clin. Investig.*, 13, 1934, 869-893.

[4] Plutchik, *op. cit.*

(1) Rises or falls starting between 34°–35° C. occurred most frequently.

(2) More falls than rises started above 35°.

(3) Approximately 70% of all changes occurred between 33° and 36°.

(4) Almost half of all rises and falls are changes of approximately 0.5°C. Only 40% of all changes were of 1° or more.

(5) The modal rate of rise is 0.6°–0.7° per min., 75% of all changes occur at a rate less than 0.7° per min.

(6) The frequencies of the various durations of rise seem to be
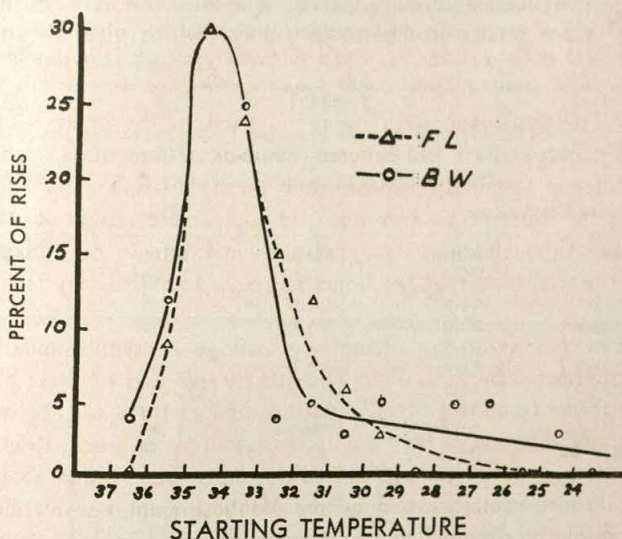
FIG. 1. RISES OF 0.5°C. OR MORE AT VARIOUS STARTING
TEMPERATURES FOR TWO Ss
(*N* for *BW* is 56; for *FL*, 33.)

bimodally distributed, the modes occurring at nearly 1 and 2 min. The most frequent duration of fall is less than 1 min. Over 90% of all temperature-changes meeting our criteria are completed in less than 4 min. It is, therefore, an extremely uncommon event for a temperature-drop to last 30 min. or more as has sometimes been reported.[5]

---

[5] E. F. Du Bois, *The Mechanism of Heat Loss and Temperature Regulation,* Stanford Univ. Press, 1937, 271-292; Bela Mittlemann and H. G. Wolff, Affective states and skin temperature: Experimental study of subjects with "cold hands" and

The difference between these and previous findings probably lies in the different criteria used to measure the termination of a temperature-drop.

As a basis for further analysis, the data of the 5 Ss who contributed the most information (in terms of the greatest number of interviews) were examined. Because of the great similarity between the two Ss most frequently interviewed, the results of their data were combined and treated *as if* they had come from one person. In other words, though the skin-temperature of *FL* varied between 29° and 35°, it is assumed that *if* it had dropped to 24° or lower it would behave in general as that of the other Ss. Upon this assumption, the data for all 5 Ss were ranked and combined. These results are shown in Table I which gives the relation

TABLE I

RELATION BETWEEN THE MEDIAN STARTING TEMPERATURE AND THE
AMOUNT, RATE, AND DURATION OF RISES OF 1.0°C. OR MORE
(Each median is based upon six scores.)

Starting Temperature		Rise		
Range (°C.)	Median (°C.)	Median (°C.)	Rate (°C./min.)	Duration (min.)
23.9–25.5	24.2	6.3	.76	9.1
26.1–27.9	26.9	3.8	.72	4.5
27.9–30.3	29.7	2.1	.76	2.2
30.6–31.9	31.4	1.9	.59	4.0
31.9–32.3	32.1	1.4	.50	2.9
32.4–33.0	32.9	1.5	.54	2.0
33.2–33.6	33.4	1.3	.56	1.9
33.6–34.8	34.2	1.1	.56	2.0

between the median starting temperature and the amount, rate, and duration of any rise of 1° or more. The total temperature-range is from 23° to 35°. It appears that the amount of rise is inversely related to the initial level at which the rise starts. The largest rises occur from the lower initial temperatures. From the same table it can be seen that the rate of rise decreases slightly with an increase in the initial temperature-level, but the change in rate is relatively small.

The duration of rise is also seen to decrease with an increase in the initial level. Table II presents the product-moment correlations for all paired sets of variables. The correlations are all high and significant at better than the 2% level.

Reynaud's syndrome, *Psychosom. Med.*, 1, 1939; and Emotions and skin temperature: Observation on patients during psychotherapeutic (psychoanalytic) interviews, *ibid.*, 5, 1943, 211-231.

TABLE II

PRODUCT-MOMENT INTERCORRELATIONS BETWEEN STARTING TEMPERATURE, AMOUNT OF RISE
OR FALL, RATE OF RISE OR FALL, AND DURATION OF RISE OR FALL

Temperature	Rises only			Falls only		
	Rise	Rate	Duration	Fall	Rate	Duration
Starting	−0.96	−0.86	−0.88	−0.13	+0.68	−0.90
Rise		+0.80	+0.96		+0.12	+0.53
Rate			+0.66			−0.29

TABLE III

RELATION BETWEEN THE MEDIAN STARTING TEMPERATURE AND THE
AMOUNT, RATE, AND DURATION OF ANY FALL
(Each median is based upon six scores.)

Starting temperature		Fall		
Range (°C.)	Median (°C.)	Median (°C.)	Rate (°C./min.)	Duration (min.)
27.8–28.8	27.8	1.2	.19	6.6
30.4–31.4	30.9	2.5	.37	6.4
31.5–32.4	32.0	1.1	.39	4.4
33.0–34.0	33.4	1.2	.27	4.0
34.3–34.8	34.4	1.2	.49	3.3
34.8–36.5	35.1	1.5	.33	4.0
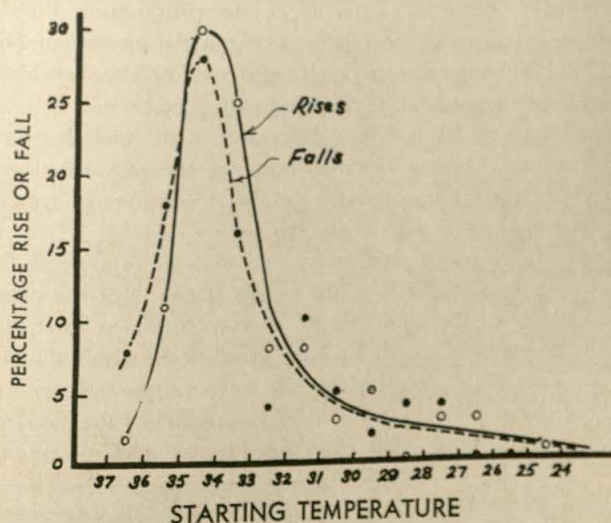
FIG. 2. RISES AND FALLS OF 0.5°C. OR MORE AT VARIOUS STARTING TEMPERATURES
(Combined data for *BW* and *FL*. N for rises in 89; for falls, 94.)

In Table III, the relation between the median starting temperature and the amount, rate, and duration of any fall of 1° or more is presented. The total temperature range is from 27.8° to 36.5°. It appears that there is no relation between the amount of fall and the initial level at which the fall starts. The correlation between these two variables is −0.13, a non-significant correlation.

The rate of fall seems, however, to be greater from higher initial levels than from lower ones. The product-moment correlation is +0.68, significant at the 1% level. The median duration of fall appears inversely related to initial level, the correlation being −0.90, which is significant at the 0.1% level. The various intercorrelations are presented in Table II.

Another point to be noted is that when a comparison was made of the rises and falls on each of the 29 records, it was found that for each the fastest rate of change of temperature from any starting point was always an increase, rather than a drop. This lends further weight to the observations already made that temperature increases seem to have different properties than temperature falls.

Discussion. The results of this study indicate that the initial level of skin-temperature existing at the time a change begins to occur is a variable influencing the amount, the rate, and the duration of change. Rises in temperature have, however, different characteristics from drops. For example, there seems to be no relation between the *amount* of fall and the initial level, although *rate* of rise is negatively correlated with starting temperature, while rate of fall is positively correlated.

These findings very likely reflect differences in the underlying physiological mechanisms, keeping in mind, however, the fact that changes in finger-temperature are not necessarily parallel to changes of temperature of other areas of the skin such as the face.[6]

Since one of the major determinants of skin-temperature is flow of blood to the surface of the skin, the results suggest that the process of forcing blood out of the capillaries by vasoconstriction to lower the skin-temperature occurs more slowly than the vasodilation which allows blood to enter the skin. This process of contraction can apparently go on independently of the already existing state of contraction of the blood vessels. Vasodilation, on the other hand, once begun is more likely to go on to the

[6] C. Sheard, M. M. D. Williams, and B. T. Horton, Skin-temperature of the extremities under various environmental and physiological conditions. In American Institute of Physics, *Temperature: Its Measurement and Control in Science and Industry,* 1941.

maximal degree possible, given the various conditions of room-temperature, humidity, sweating, etc.

It might be possible to conceive of skin-temperature as being continually driven toward a maximal value (internal temperature) by internal forces, and that any stimuli which upset the normal functioning of an individual, can only act to drive the temperature down. In this sense, drops in skin-temperature might be a relatively non-specific reaction to a wide variety of stimuli.

If, in the study of emotions, it is desired to use a more-or-less instantaneous measure of bodily change, skin-temperature is not suitable. Temperature-changes of the skin occur typically over a matter of minutes and sometimes much longer. It would be impossible to relate specific verbal content to temperature-fluctuations.

Summary and conclusions. In summary, the results of this study suggest various relationships between the variables determining skin-temperature, but they enable us to draw only limited conclusions. The situations studied have been essentially clinical rather than experimental: the content of the interviews varied from one session to another and among the Ss; and there was the additional complication that several different psychiatrists conducted the interviews.

The changes in finger-temperature of 9 Ss in a mental hospital were continually recorded during psychiatric interviews. The magnitude, rate, and duration of S's temperature-changes were related to his starting temperature, *i.e.* existing before the observed change occurred. The following major observations were made:

(1) Larger rises occur when the initial temperatures are low. The amount of fall, however, seems to be unrelated to the initial level.

(2) There is a tendency for temperature to rise at a faster rate from low than from high initial levels; whereas decreases drop faster from high than from low levels.

(3) The lower the initial temperature the longer will a rise continue, and similarly, the higher the initial temperature the longer will a fall continue, within the range studied.

(4) The fastest rate of change of skin-temperature from any starting point was always an increase rather than a drop.

It appears that rises and falls in temperature have different characteristics. Some tentative explanations were suggested, and the bearing of these findings on research in emotion was briefly indicated.

INFORMATION AND EFFECT IN INCIDENTAL LEARNING

By M. E. BITTERMAN, The Institute for Advanced Study

New evidence for the Law of Effect in human learning has recently been presented by Postman and Adams.[1] In the context of what was purported to be an experiment on *ESP*, simple English words were shown to *S*, who was asked to guess which number from *1* to *10* had been randomly assigned to each word. The responses were designated 'right' or 'wrong' in accordance with a prearranged schedule, and the consequences of this differential 'reinforcement' were examined in a second run through the list. The function of the *ESP* setting was to permit the study of 'reward' and 'punishment' independently of intention to learn in an otherwise typical Thorndikian experiment. Under intentional conditions, the authors noted, the effectiveness of reward might be attributed either to the automatic strengthening of stimulus-response connections or to the information with which *S* was provided concerning the responses to be learned, but in the *ESP* setting, they seemed willing to assume, information about the word-number pairings of one trial could not influence response on the next, and unambiguous support for the Thorndikian alternative might be found. While a previous experiment by Wallach and Henle had yielded no indication of automatic action,[2] the method employed was open to criticism. Postman and Adams, modifying the earlier procedure to minimize the possibility of interference, found more frequent repetition and better recall of rewarded as compared with punished responses. They attributed their results to the automatic action of reward.

The validity of this interpretation rests on the assumption that information played no role in the situation, and two further assumptions must be made if the role of information is to be denied. The first is that intention

[1] Leo Postman and P. A. Adams, Performance variables in the experimental analysis of the law of effect, this JOURNAL, 67, 1954, 612-631; 'Isolation' and the law of effect, *ibid.*, 68, 1955, 96-105. I am indebted to Professor Postman for information about unpublished details of his experiments and for interesting discussions of the theoretical issues.

[2] Hans Wallach and Mary Henle, An experimental analysis of the law of effect, *J. exp. Psychol.*, 28, 1941, 340-349; A further study of the function of reward, *ibid.*, 30, 1942, 147-160. Thorndike himself studied word-number associations on the pretext of interest in 'thought-transference,' but he provided no data on the effects of reward and punishment under such conditions (E. L. Thorndike, The refractory period in associative processes, *Psychol. Rev.*, 34, 1947, 234-236).

to learn was indeed entirely eliminated, and the second is that, in the absence of intent, information could not have produced any significant changes in behavior. Both are debatable, but neither will be debated here. Instead, an experiment will be reported which demonstrates the influence of information as distinct from reward on word-number associations in the *ESP* setting.

METHOD

Subjects. The Ss were 150 students drawn from elementary courses in psychology and education at San Francisco State College. They were divided into three equal groups in the following manner: The first S tested was assigned to Group I, the second to II, the third to III, the fourth to III, the fifth to II, the sixth to I, and so forth (in ABCCBA order).

Instructions. The formal instructions to the Ss were as follows (differences from group to group being set off by brackets):

Trial 1. This is an experiment in extrasensory perception—or, as it's sometimes called, mental telepathy. The procedure in such experiments is to ask a person to make guesses about events which he can't possibly perceive in the ordinary sense of the term. If the guesses are correct more often than they should be by chance alone, we may suspect that the person is acquiring information in some non-sensory fashion. Now as you probably know, there are differences of opinion on the question of whether or not there *is* such a thing as extrasensory perception, and our purpose in this experiment is to get more evidence. But before we begin, I'd like to know your opinion. I don't expect any final answer, of course, but just tell me this: If you had to guess whether or not there is such a thing as extrasensory perception, or telepathy, or thought-transference—call it what you like—would your guess be 'yes' or 'no?'

Now let's do the experiment. I have here a list of simple English words, and to each of them I've assigned a number from 1 to 10. The numbers were chosen completely at random, and I'd like to tell you how the selection was made just to give you some feeling for the situation. From a pack of ordinary playing cards, all the Jacks, Queens, and Kings were removed. That left only the Ace through Ten of each suit, of course. Then the pack was shuffled and cut to determine the number to be assigned to each word in the list. If an Ace turned up, the number assigned was 1; if a Ten turned up, the number assigned was 10; and so forth. Is that clear? Can you visualize the situation? Now what I want you to do is to try to guess the numbers that turned up. I'll read the list of words to you one by one, and as I read each word I'd like you to respond with a number from 1 to 10, whichever comes to mind first. Be spontaneous. Thinking about your performance will not help you. Just give the *first* number that comes to mind. [Group I: I won't tell you whether you're right or wrong. Group II: To let you check up on your intuitions, I'll tell you each time whether you're right or wrong. Group III: To let you check up on your intuitions, I'll tell you each time whether you're right or wrong, and if you're wrong I'll tell you what the correct number happens to be.]

Now let's practice your part of the procedure. For example, if I said [first sample word] you would say . . . ? [Group I: Fine. Group II: Then I'd tell you whether you were right or wrong. Group III: Then I'd tell you whether you were right or wrong, and if you were wrong I'd tell you the right number.] Or, I might say [second sample word] and you would say . . . ? [*The three groups are treated in the same manner as in the case of the first sample-word.*] Are there any questions? Are you ready to begin?

Trial 2. Now here is another list of words and numbers. The words are the same as those in the first list, although they appear in a different order. Numbers for

this list were chosen by the same method as those for the first list—again the pack of cards was shuffled and cut for each word, and the number assigned was the number of the card that turned up. Can you visualize the procedure? It's obvious of course that the two sets of numbers are completely independent. A given word might by chance have drawn the same number in both lists, but then again it might not, so that numbers which were right before may or may not be right again, and numbers which were wrong before may or may not be wrong again. What I'd like you to do now is to put the first list entirely out of your mind and concentrate on the second. As I read each word, give me a number from *1* to *10*, the first one that comes to mind. Be as spontaneous as possible. [Groups II and III: This time I won't tell you whether your guesses are right or wrong.] Any questions? Are you ready to begin?

The Ss were encouraged to ask questions, and every effort was made to get them to understand what was expected of them.

Procedure. The list read on Trial 1 was composed of 24 of the 25 three-letter English words used by Postman and Adams. There were two forms of the list; in one form, the words ART, MAN, SKY, BAG, and LAW (selected at random) occupied positions 5, 9, 11, 15, and 21, which were designated in advance as 'rewarded positions,' while the words GUN, WAY, EGG, TOP, and TEA (also selected at random) occupied positions 4, 10, 14, 16, and 20, which were designated as 'punished positions'; in the second form of the list, the first set of words were in punished positions and the second set in rewarded positions. One form of the list was used for half the Ss in each group, while the other form was used for the remaining Ss.

In Group I, no indication of the correctness or incorrectness of response was given to the Ss. (In an intentional learning experiment, a control procedure of this sort makes little sense to S and the results obtained are difficult to interpret, but in an ESP setting the procedure is perfectly meaningful.) In Groups II and III on the first trial, responses to words in the rewarded positions were called 'right,' while responses to words in the punished positions *and all other positions* (except the first two, which served as samples) were called 'wrong.' In Group III, all responses called 'wrong' were *corrected* as follows: For each word in the list, two numbers between 1 and 10 were selected from a table of random numbers. In the case of *JOB,* for example, the numbers were 4 and 7. If an S of Group III responded to *JOB* with any number but 4, E said, "Wrong, it's 4." In the relatively rare event that the response was 4, E said, "Wrong, it's 7." All other 'wrong' responses were 'corrected' in similar fashion. For purpose of symmetry, 'right' responses were repeated by E for Group III. If, to a word in a rewarded position, S gave the number 6, E said, "Right, it's 6."

The Ss were encouraged to respond rapidly, and for the most part they did so. On the average, the words were presented at the rate of about one every 3 or 4 sec. On Trial 2, the 20 words in positions 3-22 of the first list were presented in scrambled order. (Words in Positions 1 and 2 were used only as samples on Trial 1, while words in Positions 23 and 24 served only as buffers against a recency-effect.) The words were presented at about the same rate as before, and no indication of the correctness or incorrectness of response was given to any of the groups.

RESULTS

Three scores were computed for each *S: R,* the number of repetitions of responses to words in *rewarded* positions (of course, only the responses

of Ss in Groups II and III were actually called 'right'); P, the number of repetitions of responses to words in *punished* positions (only the responses of Ss in Groups II and III were actually called 'wrong'); C, the number of responses to words in punished positions which had previously been selected as *correct* (only the Ss of Group III actually were corrected). For example, if, on Trial 1, S responded with the number 8 to a word in a rewarded position, and if, on Trial 2, he made the same response to that word, an R was recorded—whether or not the Trial-1 response had been called 'right' (as in Groups II and III). If, on Trial 1, S responded with the number 8 to a word in a punished position, and if, on Trial 2, he made the same response to that word, a P was recorded—whether or not the response on Trial 1 had been called 'wrong' (as in Groups II and III). If the word in the punished position was JOB, for which the num-

TABLE I

MEAN NUMBER OF 'REWARDED' (R), 'PUNISHED' (P), AND 'CORRECT' (C)
RESPONSES ON TRIAL 2

Responses	Group I	Group II	Group III
R	0.88	1.12	0.84
P	0.96	0.96	0.60
C	0.34	0.44	0.62

bers 4 and 7 had been designated in advance as 'correct,' the response of 4 on Trial 2 was recorded as a C—whether or not E had announced (as in Group III) "Wrong, it's 4." If the response on Trial 1 was 4, a response on Trial 2 of 7 was recorded as a C—whether or not E had announced (as in Group III) "Wrong, it's 7." Since there were five rewarded positions and five (critical) punished positions, each of the three scores was based on five responses and therefore could take on any integral value between 0 and 5.

The mean values of R, P, and C for the three groups are given in Table I. All nine distributions were markedly skewed, with a mode at 0 in seven cases and at 1 in the remaining two. Group II showed a somewhat higher mean R than did the other two groups. The mean R of Group III was no higher than that of the control group (Group I), but its mean P was lower than that of the other two groups, and its mean C was higher. There was some tendency, it seems, for the word-number pairings defined by the consequences of response on Trial 1 to recur on Trial 2 in the data of the experimental groups. For Group II, the pairings of Trial 1—unambiguously denoted by reward alone—were reflected only in an increased R (relative to the control group).[3] For Group III, the

---

[3] This pattern—increased R without decreased P—is precisely that which appears

pairings of Trial 1—unambiguously denoted in part by reward, but principally by correction—were reflected in a shift from P to C (relative to the control group). Any tendency toward an increased R in Group III may have been counteracted (perhaps via the disruption of guessing sequences) by the predominant decline in the general level of repetition. Unfortunately, however, not too much confidence can be placed in the pattern of means on which this interpretation is based. For none of the simple scores did the difference between any two groups—evaluated by Wilcoxon's nonparametric test for unpaired replicates—approach statistical significance.[4]

Although the distributions of R, P, and C were markedly skewed, $R - P$ and $P - C$, difference-scores derived by subtraction of two frequencies, were normally distributed. Table II shows the mean difference-

TABLE II
Mean Difference-Scores and the Significance of the Deviation of Each from Zero

Group	$R-P$	t	P	$P-C$	t	P
I	−0.08	0.47	—	0.62	3.26	0.01
II	0.16	0.76	—	0.52	3.25	0.01
III	0.24	1.33	—	−0.02	0.12	—

scores for the three groups and the significance of the deviation from zero (estimated by the t-test) in each case.

Mean $R - P$ was greater in the experimental groups (II and III) than in the control (I), and greater in Group III than in Group II, but the between-groups variance did not approach significance ($F = 0.88$ with 2 and 147 df.). Nor did any mean $R - P$ significantly exceed zero (Table II). In percentage terms, the $R - P$ score was considerably smaller for Group II than for the comparable Postman-Adams groups, ($R_0 W_0$ of their first experiment and $3R, 18W$ of their second experiment). It is interest-

---

under intentional conditions (Thorndike, *Human Learning*, 1931, 38-46), and seemingly for the same reason. In a ten-choice situation, 'right' provides more information than 'wrong,' and the Ss, therefore, attend principally to responses called 'right.' Postman and I have some unpublished data which indicate that in a two-choice situation (where 'right' and 'wrong' are informationally equivalent) the decline in P is equal to the increase in R. The outcome is the same under both intentional and incidental conditions, although there is, of course, a much more marked shift on the part of intentional Ss in the direction of pairings designated (either directly by 'right' or indirectly by 'wrong') as correct on Trial 1. It is difficult to account for this difference between two- and ten-choice situations in terms of the Law of Effect.

[4] Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, American Cyanamid Company, 1949, 1-16.

ing to note, furthermore, that the largest mean $R - P$ of the present experiment (that of Group III) stemmed, not from a high R, but from a low P. Clearly, the alleged automatic strengthening effect of 'right' did not make a very impressive showing.

There was, however, a marked effect of correction. The between-groups variance in $P - C$ proved to be significant at about the 1% level of confidence ($F = 4.65$ with 2 and 147 $df.$). For Groups I and II, which were not corrected, mean $P - C$ was significantly greater than zero (Table II); that is, these Ss showed a greater tendency to repeat previously punished responses than to give the 'correct' responses. For the corrected group (III), mean $P - C$ was slightly negative; that is, the tendency to repeat punished responses was counteracted to a great extent by a tendency to respond with numbers given by E as correct on Trial 1. This outcome cannot, of course, be attributed to the action of reward, since the corrections were offered in the context of punishment. Clearly, information *per se* is established as a determinant of behavior in the *ESP* setting.

How is this result to be accounted for? One possibility is that the *ESP* procedure does not entirely eliminate intention to learn. It is certainly conceivable that not all Ss fully accept, or even fully understand, E's account of the nature of the experiment. Some may be skeptical. Others, credulous but lacking full understanding, may be led to interpret rewards and corrections as defining stable properties of the situation. The difference between the results for Group II and those for the comparable Postman-Adams groups may perhaps be explained in these terms. The procedure employed by Postman and Adams certainly left a great deal of room for suspicion. Tests for recall were made, and before being dismissed each S (although cautioned not to discuss the matter with other students) was informed that E's concern was not really with *ESP*. In the present experiment, there were no tests for recall and E's true purpose was never divulged. Furthermore, as may be appreciated by comparison of instructions, greater pains were taken here to insure that S understood what was expected of him. It may be, however, that intention to learn was not entirely eliminated even in the present experiment.

Another possibility is that incidental learning may be structured by the information provided. The knowledge that a given word has been paired with a given number, whether gained by reward or correction, may establish a closer association between the items than the knowledge that they have not been paired. In any event, the fact that information without reward may produce effects analogous to those of 'right' in the *ESP* setting

deprives the Postman-Adams data of crucial significance. The consequences of 'right' in number-guessing experiments may be uniformly understood in terms of information, but the consequences of correction cannot be understood in terms of reward. For the present, at least, the Law of Effect seems to be superfluous.

SUMMARY

As in previous experiments by Wallach and Henle and by Postman and Adams, 'automatic' effects of reward were sought in a Thorndikian number-guessing exercise conducted in an *ESP* setting which was designed to eliminate intention to learn. On the first trial, a list of words was read to S who was asked to guess which number between 1 and 10 had been randomly assigned to each. The responses of the control group (I) elicited no comment from E, but those of the two experimental groups (II and III) were designated 'right' or 'wrong.' In the case of Group III, each 'wrong' was followed by specification of the 'correct' number. On the second trial, both experimental groups showed some tendency to repeat more 'right' responses than 'wrong' responses, but in neither case was the difference statistically reliable. There was, however, a significant tendency for Group III to abandon 'wrong' responses in favor of those specified as correct by E. The fact that information as distinct from reward may influence behavior in the *ESP* setting suggests that the Law of Effect is unnecessary to account for the data of number-guessing experiments.

DEVELOPMENTAL DIFFERENCES IN THE PERCEPTION
OF CAUSALITY

By Vivian Olum, Cornell University

This paper is concerned with differences between children and adults
in the perception of causality. It employs the experimental technique de-
veloped by Michotte.[1]

Michotte made use of the fact that two colored bands, one black and one red,
painted in the form of spirals on a revolving disk and observed in cross-section
through a fixed horizontal slit, are seen as two rectangles moving horizontally. Such
a disk is illustrated in Fig. 1. The solid band represents a black object (A) and
the shaded band represents a red object (B). The dotted lines indicate the area
of the horizontal slit behind which the disk revolves counter-clockwise; O sees the
colored bands only as they pass behind this slit. Initially, B is perceived at the
center of the slit and A toward the left. As the section of band A which spirals
inward passes behind the slit, O sees A move toward B and come in contact with
it. At that point, since Band A has ceased to spiral in, O sees A stop. Almost im-
mediately, the section of Band B which spirals inward passes behind the slit, and
O sees B move towards the right and disappear behind a white cardboard, while
A remains at the center. Then A vanishes at the center. The entire process occurs
in a single revolution of the disk and hence repeats itself a large number of times.

By varying the ratio of speeds of A and B, Michotte was able to evoke a con-
tinuum of experiences from a strong causal effect of one object pushing or hitting
the other (*lancement*) to an experience of separate and unrelated movements of
the two objects. Between the clearly causal and the clearly non-causal phenomena,
Michotte found an intermediate experience in which one object, by touching the
other, triggers or releases it (*déclenchement*). This experience contains none of
the feeling of transfer of energy from one object to the other, so characteristic of
the strong causal effect. It is intermediate in the sense that the second object is
dependent on the first as initiator, although, once started, its movement becomes
autonomous.

Michotte worked primarily with adults, and there have been no de-
velopmental studies of the causal phenomenon. Nevertheless, consider-
ing the causal effect, as Michotte does, as a total configuration in time
and space made up of the separate movements of A and B, and recalling
that children have greater difficulty than adults in separating figures em-

[1] A. Michotte, *La Perception de la Causalité*, 1946, 1-296.

bedded in a larger configuration,[2] one might expect developmental differences in the perception of causality. The experiment to be reported grew out of a preliminary study with kindergarten children in which phenomena appeared (the P- and M-responses described below) that were never reported by adults and that E was unable to experience despite persistent effort.

METHOD

Observers. Two groups of 34 Os, each equally divided with regard to sex, were used. An adult group, drawn from a Cornell University summer session,
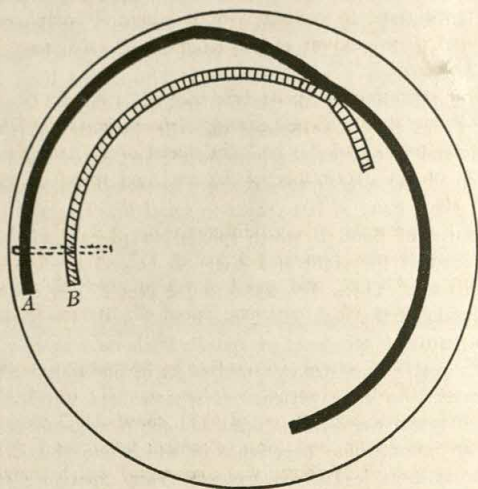
FIG. 1. DISK 3
(The dotted lines represent the slit through which the disk is seen.)

ranged in age from 17-44 yr., with a large clustering between 20 and 25, and varied widely in academic interest. The children were taken from two beginning second-grade classes.[3] They varied in age from 6.5–7.5 yr. with a mean of about 7 yr.[4] An attempt was made to choose the brightest of the available children in this

age-range with a view to equating the groups in native ability. Actually, about half of the children were considered 'bright' and the rest 'normal' or better. Rapport was generally excellent.

Stimulus-conditions. For this experiment, three disks were designed to represent three different speed-situations. All disks were of white cardboard, 50 cm. in diameter. Band *A* (black) was started 1 cm. from the edge of the disk and was 1 cm. in width. The 'chute' refers to the amount of displacement of the band towards the center per 10°, *e.g.* a 'chute' of 1 mm. on 235° means that every 10° the band has moved 1 mm. closer to the center and that the total amount of spiral occurs over 235° of angle. The 'arc' refers to the portion of *A* which does not spiral inward, but follows a circular path for a given number of degrees. Band *B* (red) started at the same radial line as *A*, but sufficiently closer to the center of the disk that it just came in contact with *A* where *A* completed its 'chute.' The 'trajectories' refer to the amount of displacement of each band from its point of origin towards the center.

Disk 1. Disk 1 represents a speed-difference of *A:B* as 1:30, for which (from Michotte's work) one should expect strong *déclenchement* (*A*, chute of 1 mm. on 235°, arc of 30°, trajectory of 2.3 cm., and speed of 1.4 cm./sec.; *B*, arc of 240°, chute of 30 mm. on 15°, trajectory of 4.5 cm., and speed of 43.2 cm./sec.; time of contact, 0.035 sec.).

Disk 2. Disk 2 represents a speed-difference of *A:B* as 1:2.33, for which one should expect a mild *déclenchement* (*A*, arc of 35°, chute of 3 mm. on 150°, arc of 100°, trajectory of 4.5 cm., and speed of 4.3 cm./sec.; *B*, arc of 190°, chute of 7 mm. on 29°, trajectory of 2 cm., and speed of 10 cm./sec.; time of contact, 0.035 sec.).

Disk 3. Disk 3 is for strong *lancement*, representing a speed-relationship of *A:B* as 1:1 (*A*, arc of 90°, chute of 9.7 mm. on 41°, arc of 135°, trajectory of 4 cm., and speed of 14 cm./sec.; *B*, arc of 135°, chute of 9.7 mm. on 26°, trajectory of 2.5 cm., and speed of 14 cm./sec.; time of contact, 0.028 sec.).

An apparatus like that of Michotte was constructed which contained three motor-driven, 20-in, aluminum disks regulated by a variable-speed drive. The absolute speed was kept constant at 2.5 sec./rev. The differences in stimulus-speed were contributed by the differences in the design of the cardboard disks which were fixed on the metal disks.

In front of each disk was a white-cardboard framework in which was cut a horizontal slit, 5 mm. high. A small *x* was drawn above each slit at the point where *A* and *B* come into contact. The length of the slits was made variable by means of sliding pieces of white cardboard. The entire apparatus was hidden from *O* by a large screen which extended about 12 ft. across the front of the apparatus and contained a center window which was opened for each observation. The room was well illuminated by overhead fluorescent lights and the apparatus-window was directly illuminated by a 200-w. lamp set about 2 ft. from the screen and to the left of the window. *O* sat in a chair placed 2 m. from this window. For each observation, *E* shifted one of the three disks into position directly behind the window. *O* could not tell exactly what was happening since the screen blocked his view, and the noise of the motor masked the noise of the moving apparatus. At each presentation, *E* opened the window and *O* saw a long, narrow opening, behind which one of the three disks revolved in a counter-clockwise direction.

Procedure. Each *O* was taken individually and asked to sit in the chair provided. Then *E* said:

This is a funny little game. Pretty soon I'm going to open this little window and you'll see two very small blocks moving, a red one and a black one. I just want you to watch where they come together and tell me what it looks like to you. What do they seem to be doing? [Motor turned on and window opened] Can you see the little blocks? What color are they? See where they come together? [Points] Right here where this little mark is. Now, watch right where they come together and tell me what's happening.

For each disk, *O* told what he saw while the phenomenon was repeating itself and *E* recorded as fully as possible what *O* said, as he said it. Only after *O* finished reporting his experience did *E* close the window and go on to the next observation. Each *O* made two observations on the most important disks, *1* and *3*, with disk *2* serving primarily as a separator. The order of presentation was in all cases *1,3,2,3,1*.

Method of analysis. The reports of the *O*s were divided into four classes as follows:

Lancement (L). This is Michotte's standard causal response. It represents all cases in which *A* pushes or hits *B* at the point of contact. (It was found necessary to judge all responses primarily at the point of contact, since children appear to have a very small *rayon d'action,*[5] and even shortening the trajectories of *B* did not eliminate its quickly acquired autonomy beyond the point of contact.)

Standard non-lancement (NL). These are the expected non-causal responses as found by Michotte when *lancement* breaks down. They include independent movement of *A* and *B*, and the intermediate *déclenchement.*

Mutual approach (M). Here the origin of activity is in *B* as well as *A*. *A* and *B* hit, bump, or push each other at the point of contact, *e.g.* "red hits black, then black hits red," or "they push each other."

Passing (P). This includes all cases in which *A* passes *B* or *B* passes *A* (as well as cases of mutual passing), *e.g.* "the black goes in front of the red," or "the red passed over the black."

RESULTS

The results of this experiment, classified into types of response, are presented in Table I. They confirm without a doubt the preliminary findings among kindergarten children. Stimulus-conditions which produce in adults responses along the normal causal continuum evoke in young children a large percentage of *P*- and *M*-responses. Specifically, 15 different children gave one or more *P*-responses while no adult ever did; and 21 children gave *M*-responses as compared with 2 adults. A straightforward probabilistic calculation shows that the difference between adults and children both on the *P*-response and on the *M*-response is significant at well beyond the 1% level. A statistical analysis of adult-child differences

[5] Michotte, *op. cit.,* 49.

PERCEPTION OF CAUSALITY 421

with respect to *L* and *NL* responses is clearly ruled out by the very large number of children for whom no *L* or *NL* assignment can be made because of their *P*- or *M*-responses.

A few of the reports (27 of the 340) contained two types, *e.g.* the

TABLE I

RESPONSES OF CHILDREN AND ADULTS TO THREE DISKS REPRESENTING
DIFFERENT EXPECTED CAUSAL RESPONSES

(Two unclassifiable child-responses and four mixed responses, three adult and one child,
showing alternation between *L* and *NL* are omitted. *Passing* includes
all responses which contain *P* in any form.)

Disks (in order of presentation)	Passing		Mutual approach		Lancement		Non-lancement	
	child	adult	child	adult	child	adult	child	adult
1	4	0	14	2	2	14	12	16
3	5	0	3	0	17	30	9	4
2	8	0	4	0	16	28	5	6
3	11	0	1	0	17	33	5	1
1	1	0	5	0	16	9	12	24

percept seemed to alternate between *P* and *L,* or between *M* and *L.* All responses appear only once in Table I. Any mixed response containing *P* is classified under *P;* all of the remaining mixed responses containing *M* are grouped under *M.* As a result, some of the *P* and some of the *M* also

TABLE II

Passing AND *Mutual-Approach* RESPONSES ANALYSED INTO SUBGROUPS

(Children's responses only are given. A corresponding table for adults would show only
two cases, both being mixed *M*- and *NL*-responses occurring on disk 1. Here
subgroups *P* and *M* refer to pure, unmixed responses.)

Disks (in order of presentation)	Subgroup				
	P	P and L	P and M	M	M and L
1	2	0	2	9	5
3	4	1	0	3	0
2	3	5	0	2	2
3	8	3	0	0	1
1	0	1	0	4*	1

* One of these responses does not quite meet the definition of *M*, since it is a mutual approach but with no form of impact at the point of contact.

contain *L* and *NL* in mixed form. Table II analyses the *P* and *M* responses into subgroups.

The passing response. In terms of the stimulus-situation, *A* is never to the right of the point of contact, nor *B* to the left of it, yet children clearly experience overlapping. The phenomenon is strikingly illustrated in some of the *P*-responses. Of the 29 *P*-responses, 7 contain reports of *A* and *B*

both moving to the right after passing, or B to the left and A to the right. In seven other cases, A and B change color as they pass, *i.e.* B turns black, or A turns red, or both. In two cases the problem presented by the passing experience is verbalized, *e.g.* "One runs past the other. The black runs past the red. It passes it somehow, over or under. The red really runs away. It looks like the black runs away too. But it still looks like one is left. It looks like the black keeps on going but it doesn't—it stops and the red keeps on going."

As can be seen in Table I, the P-responses occur primarily on Disks 2 and 3 where there is equality or relative (compared to Disk 1) equality of speeds. We know from Michotte's work that, before a causal effect can be perceived, the causing and the caused must be clearly differentiated. If we consider the movements of A and B as a configuration in time and space, it may well be that, with speeds equal, the child is unable, on the basis of color alone, to break it down into two separate movements, and that this inability interferes with causal perception here.

Mutual approach-responses. In all M-responses there is, in addition to the real movement of A from the left toward the point of contact, an induced movement of B from the right toward that point.

This phenomenon may perhaps be explained in terms of developmental differences in the perception of stroboscopic movement. Gantenbein has studied stroboscopic movement from the developmental point of view.[6] Varying the relationship between the length of exposure of the successive stimuli and the interval between them, she has found that time-relations which give an experience of successive events in adults give induced movement for children. To apply this finding to the present work, it is necessary to consider the part of B which moves to the right and disappears behind the white cardboard as the first stroboscopic stimulus and the reappearance of B at the center in a stationary form as the second stimulus. In Disk 1, it takes approximately 0.1 sec. for B to move to the right and disappear; the interval before it reappears at the center is about 0.7 sec. Extrapolating from Gantenbein's results,[7] one finds that, for a ratio of stimulus to interval of 1:7, the stimulus would have to be present for a time shorter than 0.1 sec. to produce stroboscopic movement for adults. Children, however, see stroboscopic movement up to 0.2 sec. and certainly would experience it at the exposure here employed. That this probably is the explanation of M-responses can be seen by an examination of the two other disks involved. In Disk 2, where the ratio of stimulus to interval is 1:5, a stimulus of 0.2 still would produce movement for children but not for adults. Disk 3 probably is close to the threshold even for children. A glance at the present data shows that the M-responses cluster on Disk 1, are less on Disk 2, and still

---

[6] Maria-Martha Gantenbein, Recherches sur le développement de la perception du mouvement avec l'âge (mouvement apparent, dit stroboscopique), *Arch. Psychol., Genève*, 33, 1952, 197-294.
[7] *Ibid.*, Table 4, 241.

less on Disk 3. The two cases of M-responses for adults occur on Disk 1, which is the most favorable in this respect.

It may be noted, however, that this explanation will not account for the P-phenomenon. The movement of B towards the center, in cases where B passes A and in mutual passing, may be contributed by the stroboscopic effect, but this does not explain the actual passing itself, in which the movement, real or induced, is extended beyond the limits set for it by the stimulus-situation. The distinction between the M- and P-responses in this respect is made especially obvious in terms of their occurrence on the different disks. M-responses cluster on Disk 1, which has the smallest number of P-responses, while P-responses cluster on Disk 3, the least favorable, according to the above argument, for stroboscopic induced movement.

SUMMARY

Differences between adults and children in the perception of causality were studied by Michotte's technique. Three disks of varying speed-relatións, representing different expected causal responses, were presented to 34 children (averaging 7 yr. old) and 34 adults, who were asked to describe what they saw. The main finding was that a significantly large number of children reported a *passing* phenomenon which was never experienced by adults. If the movements involved are considered as a configuration in time and space, the child's relative inability to break down configurations may be responsible for the phenomenon. A secondary finding was that children gave *mutual-approach* responses significantly more often than adults. A possible explanation of this effect is afforded by developmental differences in the perception of stroboscopic movement.

THE RELATIONSHIP BETWEEN ASSOCIATIVE FREQUENCY, ASSOCIATIVE ABILITY AND PAIRED-ASSOCIATE LEARNING

By George Mandler and Janellen Huttenlocher, Harvard University

One of the most widely accepted generalizations about human verbal learning concerns the relationship between meaningfulness of material and facilitation in the speed of learning. The determination of the associative value of nonsense syllables by Glaze and others was designed to provide the experimenter with an appropriate control for variations in meaningfulness.[1] Subsequently, Melton and McGeoch showed that lists of nonsense-syllables are learned differentially as a function of the associative value of the nonsense-syllables which make up these lists.[2] Guilford also showed that lists of nonsense-syllables require more trials to learn than lists composed of three-letter words.[3] Sheffield obtained similar results showing the advantage of meaningful material (adjectives) over less meaningful material (nonsense-syllables).[4]

Despite this general agreement concerning the effect of meaningfulness on learning, the problem has received little theoretical attention. Recently, Noble has postulated that 'meanings' of a stimulus-increase as a function of the number of S-multiple R connections.[5] He has also developed an index of meaningfulness (m) which is defined as the mean-frequency of continued associations evoked by a stimulus in 60 sec. Subsequent work has shown that this measure m, determined for 96 dissyllables, predicts difficulty of learning and errors in position in serial, verbal learning.[6]

Some theoretical considerations by Mandler lead to similar predictions.[7] This point of view suggests that in adult learning new associations are mediated by previously learned associations to the stimulus- and response-terms. From this it can be deduced that the more such associations are elicited by the stimulus- and response-members in a paired-associate task, the greater will be the probability that a particular combination of a stimulus-association and a response-association can be found to provide a mediating bridge or a 'meaningful' link between the two. Two further deductions from this theoretical position concern individual differences. In the first place, we would predict that the more associations an individual has

[1] J. A. Glaze, The association value of nonsense syllables, *J. genet. Psychol.*, 35, 1928, 255-269.

[2] A. W. Melton, A comparative study of the materials employed in experimental investigations of memory, unpublished Master's thesis, Yale Univer., 1929. J. A. McGeoch, The influence of associative value upon the difficulty of nonsense-syllable lists, *J. genet. Psychol.*, 37, 1930, 421-426.

[3] J. P. Guilford, *Laboratory Studies in Psychology*, 1934.

[4] F. D. Sheffield, The role of meaningfulness of stimulus and response in verbal learning, unpublished Doctoral dissertation, Yale Univer., 1946.

[5] C. E. Noble, An analysis of meaning, *Psychol. Rev.*, 59, 1952, 421-430.

[6] Noble, The role of stimulus-meaning (m) in serial verbal learning, *J. exp. Psychol.*, 43, 1952, 437-446.

[7] George Mandler, Response factors in human learning, *Psychol. Rev.*, 61, 1954, 235-244.

available, *i.e.* the more associations he produces to verbal stimuli in general, the more quickly will he learn new associations between two nonsense-syllables. Another prediction is based on the assumption that speed of learning will be impaired if *S* responds with an equal frequency of association to *all* syllables. In that case, no particular syllable pair in a list would have any advantage, and interference between pairs would be at a maximum. This prediction states that there will be a positive correlation between speed of learning and the degree to which *Ss* show *differential* associative behavior in response to different nonsense syllables.

This study presents data relevant to these three predictions. In a previous study data have been presented which determined associative frequency (*f*-value) for 100 nonsense-syllables.[8] This *f*-value was the mean number of associations elicited by nonsense-syllable stimuli in a 30-sec. interval. We have used this measure as an indicator of the tendency of stimulus- and response-members to elicit previously acquired associations. We wish to see whether speed of learning is related to this *f*-value. In addition to the paired-associate learning task, *Ss* were given an association-task in which their individual associative frequency to nonsense-syllables (different from those used in the learning task) was determined. From this, we hoped to test the predictions based on individual differences.

METHOD

Materials and procedure. Two different lists (A and B) of eight paired nonsense-syllables were constructed for the learning task. The syllables were selected from the list of 100 consonant-vowel-consonant syllables whose associative frequency-values *(f)* have been published. The syllables were selected at random from eight

TABLE I

NONSENSE-SYLLABLES WITH THEIR RESPECTIVE *f*-VALUES AND MEANS AND VARIANCES OF SCORES IN LEARNING

	List A					List B			
Syllable pairs	f-values		Mean acquisition pairs	σ^2	Syllable pairs	f-values		Mean acquisition trial	σ^2
	stimulus	response				stimulus	response		
XOK-ZUV	2.9	3.1	21.00	86.00	VOJ-KEF	3.1	3.5	16.15	11.29
QIJ-XUR	3.4	3.4	17.60	98.88	PUQ-XIV	3.6	3.1	12.20	80.06
ZEG-KOV	3.5	3.5	20.20	130.69	GUV-SIJ	3.6	3.6	17.20	111.75
WUF-NAZ	3.9	4.0	16.55	78.79	DIF-GAN	3.8	3.9	9.45	19.73
HAV-BIP	4.1	4.3	11.70	89.80	BIZ-VEW	4.2	4.1	13.65	68.34
SIK-FEY	4.4	4.5	10.25	37.98	YAP-CID	4.6	4.5	12.35	39.82
LIR-WAT	4.7	4.8	16.35	93.71	SAX-LIM	4.7	4.7	13.20	79.33
VIN-PAS	5.0	5.3	7.20	19.22	LAT-MON	5.3	5.1	7.60	10.78

equal intervals of the distribution of *f*-values. Thus, for List A, one stimulus-response *pair* was selected from the first interval, one pair from the second, and so on. The selection was restricted by the further requirements that the same letter did not appear in both the stimulus- and response-terms, any initial letter appeared only once in a set of stimulus- or response-terms within lists, and no pair of letters appeared in more than one syllable. List B was constructed similarly, except that Intervals 1 and 2 had to be combined to produce enough syllables to satisfy all conditions. Table I shows the syllables used in both lists with their respective *f*-values.

[8] Mandler, Associative frequency and associative prepotency as response measures of nonsense syllables, this JOURNAL, 68, 1955, 662-665.

The lists were presented in a Hull-type memory-drum. The stimulus-member of the pair was exposed for 1.5 sec., followed by a 0.5-sec. interval during which both shutters were closed, and a 1.5-sec. exposure of the response-member, with another 0.5-sec. interval during which both shutters were closed.

In the association-task *S* was seated at a table and successively presented with 20 nonsense-syllables printed on 3×5-in. filing cards. Each syllable was presented for 30 sec., during which time *S* wrote down his associations. Previous tests have shown that this time-interval is sufficient to elicit *S*'s readily available associations. The 20 syllables represented the full range of *f*-values. No syllable appeared in both the learning and the association-task for any particular *S*. The two sets of 20 syllables used for the association-task had 12 syllables in common. One association-list was used with each of the learning-lists.

Order of presentation of the learning and the association-task were balanced among *S*s reacting to each list. Half of the 20 *S*s who learned a particular list were given the association-task first, and the learning-task second (order *AL*); the other half of the *S*s were given the tasks in reversed order (order *LA*).

Subjects and instructions. The *S*s were 40 paid volunteer graduate and undergraduate students who had not previously participated in any similar experiment. Prior to the association-task, all *S*s were instructed to give only associations to the stimulus-syllable, and were cautioned against giving chain associations, *i.e.* responses that were not directly evoked by the nonsense-syllable. No special instructions were given for the learning-task, and the connection between the two tasks was not explained to the *S*s.

RESULTS AND DISCUSSION

The learning-measure used for any syllable pair was the number of the second of two successive trials in which the correct responses were given. Table I shows the means and variances of this measure for both lists. Since these measures showed marked heterogeneity of variance both within pairs and within *S*s, a logarithmic transformation was used for purposes of statistical analysis. Table II shows the results of the analyses of variance of the learning-measure for both lists. The only significant source of variation derives from differences between syllable pairs, *i.e.* differences in associative frequency. While the variance contribution from linear regression and the deviation from linear regression are both significant, it should be noted that for List A linear regression accounts for significantly greater part of the variance between pairs than does the remaining variance, while for List B this value reaches the 10% level of significance.

The transformed data were further tested for the relationship between associative frequency and speed of learning by correlating the mean associative frequency-value and the mean learning-measure for the eight pairs. For List A, the product moment correlation was -0.82 ($p < 0.01$), for List B, it was -0.62 ($p < 0.05$). Thus, for both samples of *S*s and lists, we find a significant relationship indicating faster acquisition as a function of increased associative frequency-values. Inspection of Table I shows that the deviation from a linear relationship is not a simple curvilinear one, but rather shows a delayed acquisition of the second highest pair in both lists (LIR-WAT and SAX-LIM) and a surprisingly fast acquisition of the second lowest pair in List B (PUQ-XIV). While we can offer no ready explanation

for the former deviation, it was noted that in the case of the PUQ-XIV association, many Ss used the Roman numeral fourteen as a mediating link. This possibility, which we did not foresee, might have placed this pair perceptually outside of the population of nonsense-syllable pairs and thus facilitated acquisition.

Two procedures were used to test the hypotheses related to individual differences. The first used the sum of the associations each S produced to the 20 association syllables. This measure was correlated with the sum of each S's learning scores for all eight syllable pairs. While in the predicted direction, the correlations of -0.35 and -0.08 for Lists A and B are not significantly different from zero. The second procedure was designed to measure the S's *differential* associative frequency and to relate it to speed of learning. For this purpose we obtained S's number of associations

TABLE II

Analyses of Variance of Transformed Learning Scores

Source of variance	df	List A		List B	
		Mean square	f	Mean square	f
Order of presentation	1	.017	< 1	.006	< 1
Error (between Ss)	18	.221	—	.231	—
Syllable pairs	7	.549	11.20†	.241	12.05†
Linear regression	1	2.698	55.06†	.652	32.60†
Deviation from linearity	6	.191	3.90*	.172	8.60†
Order × pairs	7	.031	< 1	.039	1.45
Error (within Ss)	126	.049	—	.020	—

* Significant at the 1% level.
† Significant at the $\frac{1}{10}$% level.
Note that f-values for linearity of regression tested against remaining variance (deviation) are 14.13 (p=0.01) for List A and 3.79 (p=0.10) for List B.

to the upper 10 and the lower 10 nonsense syllables arranged in descending order of the previously determined associative frequency. A partial correlation was obtained between speed of learning and the number of associations given to the 10 high syllables, holding number of association to the 10 low syllables constant. These partial correlations were -0.64 (p < 0.01) for List A, and -0.57 (p < 0.01) for List B. One way of conceptualizing these findings is that the number of associations given to the lower ten syllables represent a general level of associative fluency, and that there is a significant relationship between number of associations elicited and speed of learning when this presumably interfering fluency factor is held constant.

It will be noted that List B shows consistently lower values for the predicted relationships. To explore this difference, an analysis of variance of all 40 Ss' associative scores was performed, which failed to show any significant differences associated with either lists or order of presentation. The associative scores of the Ss who learned List B showed, however, a significantly greater variability than those of the Ss in List A ($f = 2.54$, $p < 0.05$). Since Ss were assigned randomly to the counterbalanced conditions, but not to the two lists, the greater variability of the Ss on List B may account for the lower relationships found.

Our results substantiate the general hypothesis that associative frequency is an

important determiner in paired-associate learning. Kimble and Dufort, have found a similar relationship between paired-associate learning and Noble's m value.[9] It should be pointed out, however, that we have shown previously that our f-measure is related ($r = 0.65$) to association values similar to Glaze's, which presumably might predict paired-associate learning equally well. We would suggest, however, that these findings are conceptually more fruitful when a theoretically derived measure is employed.

The failure to find significant support for our main individual difference hypothesis may be a function of the homogeneity of our student sample. The finding that a measure derived from an S's associative behavior is related significantly to learning ability is reassuring, particularly in the light of past unsuccessful attempts to relate individual difference variables to nonsense-syllable learning.

SUMMARY

The purpose of this study was to test the hypothesis that the more associations each member of a stimulus-response pair of nonsense syllable elicits, the more easily can they be associated with each other. Further predictions concerned the relation between individual differences in frequency of associations and speed of paired-associate learning.

Two groups of Ss each learned a different paired-associate list. The eight stimulus-response pairs of each list represented eight different levels of associative frequency (f), $i.e.$ increasing tendencies for the nonsense-syllables to elicit continued associations. The Ss also gave 30-sec. continued associations to 20 syllables different from those used in the learning task.

The main results were:

(1) A significant correlation between speed of acquisition of a nonsense-syllable pair and the mean associative frequency previously measured for its two components.

(2) A positive, but non-significant, relationship between speed of learning and number of associations given by S in the association-task.

(3) Confirmation of the hypothesis that there is a positive correlation between individual differences in associative frequency and speed of learning when associative fluency or differential associative behavior are taken into account.

[9] G. A. Kimble and R. H. Dufort, Meaningfulness and isolation as factors in verbal learning, $J. exp. Psychol.$, 50, 1955, 361-368.

THE EXTENSION OF MARBE'S LAW TO THE RECALL OF STIMULUS-WORDS

By W. A. Bousfield, B. H. Cohen, and Joan G. Silva, University of Connecticut

In one of the early studies of free association Thumb and Marbe reported an inverse curvilinear relationship between individual reaction-times for freely associated responses and the frequency of occurrence of these responses derived by pooling group data.[1] This negative relationship between frequency of responses and the speed of their emission has come to be known as Marbe's Law. Woodworth has reviewed evidence of its confirmation by several investigators.[2] The present study reports findings indicating that Marbe's law may be extended to the recall of the words of stimulus-lists. In general terms our procedure was that of presenting the words of stimulus-lists in serial order for learning to groups of undergraduate student Ss who could be assumed to have fairly similar cultural and educational backgrounds. These Ss were instructed to write as many of the stimulus-words as they could recall in the order of their occurrence in memory. Within the framework of these operations we undertook to test the following hypothesis: *The rank order of recall of the words of a stimulus-list by individual Ss should be a negative function of the frequency of recall of the words by the group.*

In accounting for this deduction, we assume that both Marbe's law and our extension of it should apply when the Ss comprising the group have more or less similar habit-strengths for the individual words comprising the stimulus-list. Such similarities should be a function of the degree of homogeneity in the social and educational backgrounds of the Ss. In the so-called free association experiment, the S responds to a stimulus-word by drawing from a pool of competing associates having the common characteristic of being related to the stimulus-word. Under these conditions the individual S produces the associate having the greatest strength. This strength, in turn, may be measured by reaction-time. Within the conditions of the present study, the Ss were required during recall also to draw upon a pool of associates. It would follow that associates having the greatest strength should be produced first. The order of recall of the subsequent words, moreover, should be a function of their strengths. It should be noted here that in treating our data we split the recall sequences of the Ss into decile intervals. This method is consistent with our theorizing especially if we can accept the reasonable assumption that decile rank and response latency are directly related by some monotonic function.

PROCEDURE

Stimulus-words. We used five different lists of stimulus-words classified in three groups as follows: (a) List I comprised 40 unrelated nouns chosen so as to attain

[1] A. Thumb and Karl Marbe, *Experimentelle Untersuchungen über die psychologischen Grundlagen der sprachlichen Analogiebildung*, 1901. For this reference the authors have relied upon the citations of C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 722-723, and on Woodworth, *op. cit.*, 360-362.

[2] R. S. Woodworth, *Experimental Psychology*, 1938, 360-362.

a somewhat uniform sampling of the range of Thorndike-Lorge general counts of frequencies-of-usage.[3] We prepared two separate randomizations of this list, here designated as List I-ₐ and List I-ᵦ. (b) List II-c comprised 20 words in each of two categories; namely *animals* and *musical instruments*. List III-c comprised 10 words in each of four categories; namely, *animals, clothing, musical instruments,* and *weapons.* List IV-c comprised five words in each of eight categories; namely, *animals, birds, clothing, countries, elements, musical instruments, ships,* and *weapons.* The words of these three lists were also chosen by sampling from the Thorndike-Lorge general counts. (c) List V was a 60-word list of nouns comprising 15 in each of four categories; namely, *animals, names, professions,* and *vegetables.* The words of this list had Thorndike-Lorge general counts of frequency-of-usage within the range of one to 19 per million. The mean frequency-of-usage for each of the four categories was 7.33. We prepared this list for studying the effects of multiple presentations of the stimulus-list on learning. In this undertaking we wished to minimize the effect of serial position when the list was presented more than once to the Ss. This was accomplished by preparing five different randomizations so as to avoid the repetition of serial order in the stimulus-words. We shall here deal with the results obtained by presenting the list once to one group of Ss, three times to a second group, and five times to a third group. For convenient reference we have labelled List V in terms of the number of times it was presented, *i.e.* reinforced, for the corresponding groups as follows: V-1 reinf., V-3 reinf., and V-5 reinf.[4]

Apparatus. The words of all the lists were copied on glass slides. To expose the words on a screen we used a Keystone overhead projector equipped with a mask having a rectangular opening. The E exposed the words singly and at 3-sec. intervals by moving the slides over the opening in the mask in time with the flash of a small light mounted on the projector. Data sheets 8.5 × 11 in. in size were prepared for use by the Ss in writing the words they were able to recall.

Procedure. A uniform procedure was followed in presenting the words of the stimulus-lists to groups of Ss for learning. The experiment was conducted in the lecture or laboratory room where the Ss regularly met for class. Having requested the coöperation of the Ss, E read a prepared statement of instructions indicating that a list of words would be exposed one at a time on the screen. Immediately after the completion of the presentation of the last word, at a signal given by E, the Ss were to write the words they were able to recall on the data-sheets and to write them in the order of their occurrence in memory. The E then presented the stimulus-words and 3-sec. after the completion of the presentation gave the signal to the Ss to start writing the words they were able to remember. A total of 10 min. was allowed for recall. Using this procedure, Lists I-ₐ, I-ᵦ, II-c, III-c, and IV-c were presented once to the respective groups of Ss. In the case of List V, however, where we were concerned with the effects of multiple reinforcements, the words were presented once to one group of Ss, three times to a second group, and five times to a third group.

---

[3] E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words,* 1944.

[4] Lists II-c, III-c, IV-c, and V were prepared for use in studies of associative clustering. We are here reporting our recent analysis of data obtained from the use of these lists in earlier experiments.

Subjects. Separate groups of Ss were used for each of the stimulus-lists as follows: List I-a, 73 Ss; List I-b, 52 Ss; List II-c, 41 Ss; List III-c, 35 Ss; List IV-c, 53 Ss; List V-1 reinf., 49 Ss; List V-3 reinf., 46 Ss; List V-5 reinf., 47 Ss.

RESULTS

We undertook several different analyses of the data. The following was chosen for presentation here because it appeared most effectively to reveal the essential nature of the results. For each group of Ss a master data-sheet was prepared with all the words of the stimulus-list written in a column at the left. The recall-sequences for each of the Ss were divided by the Vincent method into deciles that were numbered in order from one to 10.[5] The first column to the right of the words on the master-sheet was used for entering the data for the first S, the second column for the second S, and so on for all the Ss in the group. In recording the data for an individual S we entered on the master-sheet the decile of each stimulus-word as it appeared in the recall-sequence. This entry for each recalled stimulus-word was made opposite the same word on the master-sheet.

Having recorded the deciles of all the stimulus-words recalled by all the Ss of the group, we obtained the frequency of recall for each stimulus-word by adding the number of entries made in its row on the master-sheet. The mean decile in which each word appeared in recall was obtained by averaging the decile numbers in its row. The frequency of recall for each stimulus-word and the mean decile in which it appeared in recall are our two essential indices for summarizing the data. To test our experimental hypothesis stating the expected negative relationship between rank order of recall of the words by the individual Ss and the frequency of recall of the words by the group we computed the Pearson r between these variables for each of our eight groups of Ss. We also derived the regression equation for each correlation and applied the t-test to determine the significance of the difference between the slope of the regression line and a slope of zero.

Table I summarizes the results of these analyses. We may see from this table that all the correlations are negative and hence in the direction predicted by the experimental hypothesis. The regression equations are in the form of $y = k + mx$, with k as the intercept of the regression line on the y-axis and m as its slope.

For a further analysis of the data we undertook to test the null hypothesis that the variability of the individual regression lines represents random sampling from a homogeneous population. The pooled sums of squares from the individual regressions were used to compute an average regression equation with appropriate sums of squares. Using the treatment of Snedecor, the F for the null hypothesis of random sampling from a homogeneous population was found to be 0.67.[6] Thus, we have no reason for rejecting the null hypothesis. The average regression equation and average correlation are shown in the last line of Table I. From this further analysis we may conclude that the regression equation derived from the pooled data may be regarded as having sufficient generality to describe the results obtained from all eight groups of our Ss. Thus the predicted relationship emerges under three

[5] S. B. Vincent, The function of the vibrissae in the behavior of the white rat, *Ani. Behav. Monogr.*, 1, 1912 (No. 5), 16-17.
[6] G. W. Snedecor, *Statistical Methods*, 1950, 325-327.

conditions of use of stimulus-word lists: (a) unrelated words; (b) words in categories; (c) multiple reinforcements of words in categories. It should be noted that Condition (c) refers to the use of List V—3 and List V—5. In these cases

TABLE I

CORRELATIONS BETWEEN FREQUENCIES OF RECALL OF STIMULUS-WORDS BY EXPERIMENTAL GROUPS AND MEAN DECILES OF THEIR OCCURRENCE IN RECALL

Stimulus-Word	No. Ss	r	Regression equation	t^*
Unrelated Words				
List of I_a Randomization 1	73	$-.796$	$y = 7.94 - .084x$	8.14
List I_b Randomization 2	52	$-.454$	$y = 7.18 - .058x$	3.14
Words in categories				
List II_C, 2 categories	41	$-.418$	$y = 7.08 - .061x$	2.87
List III_C, 4 categories	35	$-.251$	$y = 6.27 - .042x$	1.59
List IV_C, 8 categories	53	$-.642$	$y = 7.62 - .073x$	5.18
Words in 4 categories, varying reinforcement				
List V-1 reinf.	49	$-.384$	$y = 6.58 - .051x$	3.18
List V-3 reinf.	46	$-.515$	$y = 7.35 - .070x$	4.59
List V-5 reinf.	47	$-.473$	$y = 7.45 - .067x$	4.09
Pooled Results for all Lists	396	$-.541$	$y = 7.27 - .068x$	—

* All ts are significant at the 1% level or better except for the case of List III_C.

we used separate randomizations for each presentation of the list in order to minimize serial position effects.

DISCUSSION

Our experimental hypothesis is clearly noncommittal regarding the specific sources of the habit-strengths of the stimulus-words at the time of their recall. We are inclined, however, to discount the influence of serial position effects. Our treatment of the data for the two groups for whom this influence was minimized revealed no unique characteristics. It appears to us that both Marbe's law and our confirmation of one of its corollaries demonstrates the potency of cultural and educational influences in building up the differing strengths of verbal habits. It is beyond the scope of this paper to undertake a review of the evidence demonstrating the potency of culture in conditioning verbal behavior. We regard it as relevant, however, to mention a study rather closely related to the present one. Bousfield and Barclay reported evidence indicating a negative relationship between the rank order of occurrence of restricted associative responses given by individual Ss and the frequency of occurrence of the responses for the group.[7] Restricted associative responses were obtained on the basis of instructing groups of Ss to list items belonging in the same category. Specifically, they analyzed repeated associations that their Ss produced when asked to make lists of *birds, carpenter's tools,* and *celestial bodies.* Their analysis showed the same type of relationship between group-frequency and rank-order of individual response as was demonstrated in the present study. It should be

---

[7] W. A. Bousfield and W. D. Barclay, The relationship between order and frequency of occurrence of restricted associated responses, *J. exp. Psychol.,* 40, 1950, 643-647.

noted here that Johnson has pointed out that Marbe's law may be extended to apply to the continuous production of words and that he used the Bousfield and Barclay study as an illustration of this extension.[8] If our reasoning is correct, we may speculate that the basic phenomenon with which we are dealing should emerge whenever groups of Ss having generally similar cultural backgrounds are required to produce verbal items from a similar pool of associates. Beyond this it is tempting to deduce that with the proper choice of stimulus-words the experimental method we have employed may be used effectively as a device for appraising the cultural homogeneity of the verbal learning of groups of Ss.

SUMMARY

The purpose of this study was to test the following hypothesis which the authors interpret as a corollary of Marbe's law: *The rank order of recall of the words of a stimulus-list by individual Ss should be a negative function of the frequency of the recall of the words by the group.* It is assumed that this hypothesis is applicable when the Ss constituting the group show an optimal degree of cultural homogeneity. The method was that of presenting stimulus-lists of unrelated words and of words in categories to eight different groups of Ss for learning. The Ss were instructed to write the words they were able to recall in the order of their occurrence in memory. Support for the experimental hypothesis was indicated by the negative correlations existing between rank order of recall of the stimulus-words by individual Ss and the frequency with which the words were recalled by the groups.

[8] D. M. Johnson, *The Psychology of Thought and Judgment,* 1955, 186-187.

ACCURACY OF A SIMPLE POSITIONING RESPONSE WITH VARIATION IN THE NUMBER OF TRIALS BY WHICH KNOWLEDGE OF RESULTS IS DELAYED

By INA McD. BILODEAU, Lackland Air Force Base

That relatively brief temporal delay of knowledge of results *(KR)* is an effective variable in learning may be questioned if simple, discrete motor responses are considered. It is difficult, at least, to locate effective manipulations of delay in the literature, the findings typically being inconclusive.[1] In contrast, small lags in transmitting information greatly influence continuous tracking performance.[2] Among the differences between feedback-delays in, say, a simple positioning task and a tracking task, one seems particularly important: that *KR* is interpolated before a second positioning response is made, while *S* must continue responding in a tracking task before he has feedback for previous manipulations. The effects of lag in a tracking task may therefore depend on the correction of on-going behavior by *S* on the basis of a signal appropriate to behavior earlier in the sequence.

The importance of this difference can be tested by so arranging delay in the simple positioning task that it approaches the transmission type of control-lag. The present technique, delaying every *KR* for a number of trials, makes for such an arrangement, requiring *S* to respond without benefit of *KR* about his last response-error. Lorge and Thorndike,[3] using the trial-delay technique, found a sizable effect—indeed, they found no evidence of improvement when another response of the same kind was interpolated between a response and the *KR* to be associated with it.

The experiments to be reported repeat the trial-delay technique on a simple lever-positioning task, the number of trials over which the *KR* is delayed being the variable. The device selected has shown no change in behavior attributable to variation in temporal delay alone.[4]

Subjects. A total of 336 airmen, basic-trainees at Lackland Air Force Base, served as *Ss*. The *Ss* were assigned to treatment-groups as they appeared, according to successive, unsystematic schedules using each group once. Experiments I and II were run successively.

Apparatus. The apparatus was the Manual Lever previously described in detail,[5]

[1] Irving Lorge and E. L. Thorndike, The influence of delay in the after-effect of a connection, *J. exp. Psychol.,* 18, 1935, 186-194; L. T. Alexander, Knowledge of results and the temporal gradient of reinforcement, Unpublished Ph.D. Dissertation, Ohio State University, 1950.

[2] M. J. Warrick, Effect of transmission-type control lags on tracking accuracy, USAF Technical Report No. 5916, Wright-Patterson Air Force Base, 1949.

[3] Lorge and Thorndike, *op. cit.,* 186-194.

[4] E. A. Bilodeau, unpublished data.

[5] E. A. Bilodeau and T. G. Ferguson, A device for presenting knowledge of results as a variable function of the magnitude of the response, this JOURNAL, 66, 1953, 483-487.

used with a 14.5-in. lever-arm and a required displacement-force of 20 lb. The device consists essentially of a lever-arm shielded from *S*'s view, the panel from which *E* reads the magnitude of lever-travel and controls *KR*, and *S*'s *KR*-display. The *S*'s *KR*-display has a 100-unit scale (each unit equal to 0.746° of arc of lever-travel) with a pointer to indicate the score, *i.e.* the amplitude of displacement. The lever automatically falls back to zero after *S* makes a response and releases the lever.

The task was to learn to make the lever-displacement corresponding to a score of 45 (33.57° of arc). One pull or displacement constituted a trial. Scores were reported without distortion other than that introduced by rounding to the nearest whole number.

The trial-cycle was 20 sec. in duration, allowing 5 sec. between ready and pull signals, time as required for the response (about 4 sec.), and approximately 11 sec. between response-completion and the next ready-signal; *KR* was interpolated about 5 sec. before the ready-signal.

Procedure. Groups of 48 *S*s were differentiated by the number of trials over which *KR* was delayed. The delays were 0, 1, 2, and 3 trials in Experiment I, and 0, 2, and 5 trials in Experiment II. *E* gave all *S*s full information about their delay-treatments and carefully questioned them to assure their understanding.

At the start of the series of responses, the trial-delay meant that *KR* was not given on some trials and that the number of such trials varied with the delay. That is, for example, a 1-trial delay meant that *S* received the score for his first response only after completing his second response, a 5-trial delay meant that the score for the first response was not given until the sixth response was made. Once *S* received his first *KR*, of course, he received *KR* on all succeeding trials, but every *KR* was delayed by the given number of trials.

Within an experiment, the trials held constant in number from group to group were *KR*-trials, defined as trials preceded by *KR* (*i.e.* numbering of *KR*-trials began with the first trial on which *S* was not responding blindly). There were 16 such *KR*-trials in Experiment I, necessitating totals of 17, 18, 19, and 20 responses, respectively, for *S*s with delays of 0, 1, 2, and 3 trials; 0-delay *S*s were given in addition one practice trial without *KR* before the start of their regular 17-trial series. The 30 *KR*-trials of Experiment II required totals of 31, 33, and 36 trials, respectively, for delays of 0, 2, and 5 trials.

After completing the 16 *KR* trials of Experiment I, all *S*s had a 1-min. rest before a 4-trial test under the 0-delay treatment. The test was omitted in Experiment II.

Results. Data presented in this section have been converted from the arbitrary score-scale units to positioning error in degrees of arc. The trials to be considered are *KR*-trials; even on the first trial, therefore, there can be differences between groups, since the first *KR* refers to a trial progressively more remote from *KR*-trial 1 as the trial-delay increases. Blind trials, before any feedback was given, have been omitted to separate the effects of delay and number of *KR*s. In general, error did not change until *KR* was introduced.

In Fig. 1, mean absolute error is plotted against successive pairs of *KR*-trials. All groups give evidence of learning, in contrast to the finding of Lorge and Thorndike with a 1-trial delay.[6] In both experiments, nonetheless, the 0-delay group has the

[6] Lorge and Thorndike, *op. cit.*, 186-194.

smallest mean errors and the groups with trial-delay differ from each other. In Experiment I, error increases with increasing trial-delay, except for the over-all lack of difference between 2- and 3-trial delays. In Experiment II, where the range of delays and the differences between adjacent delays were greater, the three curves are clearly separated.

Delaying *KR* over interpolated trials thus has an effect similar to that of lag in a tracking task, increasing response-error relative not only to a 0-delay treatment, but also relative to lesser delays. These effects of trial-delay are consistent with the
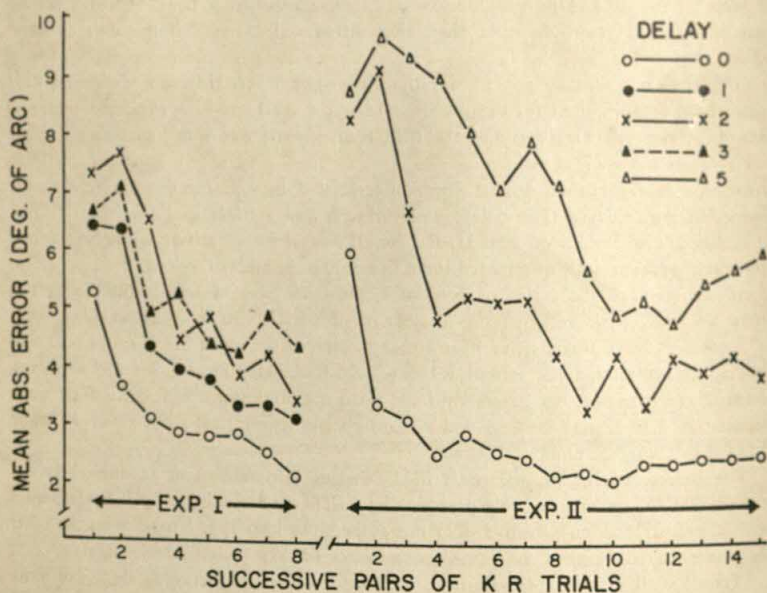
FIG. 1. MEAN ABSOLUTE ERROR IN LEVER-PLACEMENT FOR SUCCESSIVE PAIRS OF *KR*-TRIALS

(Each data-point is the mean of the mean errors of the pair of trials)

position that the 0-delay *KR* serves a directive function, telling *S* how his response differs from the required response and indicating the magnitude and direction of response-change needed to make the next response more accurate than the last.[7] In the simplified 0-delay case, for example, *S* is told that his score was 50 on his last trial; this means that *S* pulled beyond 45, and the *KR* can be considered a signal to pull less far next time. Applying the 0-delay reasoning to the trial-delay case, we see that the *KR* no longer tells *S* how his last response can be improved. The delayed *KR* is an incomplete error-signal, lacking information about response-changes made between the trial reported and the last trial. In using the *KR* to improve the accuracy of his next response, *S* must take intervening response changes into account —and a larger number of changes, the greater the trial-delay. Since the effects of

---

[7] Bilodeau, Speed of acquiring a simple motor response as a function of the systematic transformation of knowledge of results, this JOURNAL, 66, 1953, 409-420.

any intervening changes have not been reported, there may be sizable discrepancies between the amount by which S thinks he has changed his response and the amount by which he actually has changed it. With trial-delay, in brief, S must estimate (without feedback) how his last response differs from the response to which KR applies, as well as take into account the error signalled by the KR. The trial-delay case, therefore, suffers the simple estimation- and response-errors of the 0-delay case and introduces such error-sources as forgetting, interference, incomplete information, and the compounding of simple errors.

Most of the delayed groups have a rise in mean error from the first to the second pair of KR trials (see Fig. 1). The phenomenon was evident in the data of most of the Ss, and apparently depended on S's reversing the direction of his initial error and over-compensating in the magnitude of his correction. Since most Ss consistently overshot or undershot in the blind trials, the early KRs reported errors in one direction only (for the given S). Trial-delay may, therefore, lead to the additional error-tendency of confusing the correction required at the moment with the correction relevant to some earlier trial. At least S may give undue weight to the immediately preceding KR in a series suggesting a constant error.

A Groups $\times$ Trials—or Lindquist's Type I*—analysis of variance was computed for each experiment. In Experiment I, since the variances of the absolute errors decreased over the first six or seven KR-trials and were fairly stable thereafter, only KR-trials 6-16 were treated. The F-test of the Groups main effect ($F = 4.00$, $df. = 3$ and 188) was significant at the 1% level and that of the Groups $\times$ Trials interaction ($F = 1.70$, $df. = 30$ and 1880) just reached the 1% level. In Experiment II (KR-trials 7-30), tests of both Groups and Groups $\times$ Trials effects were significant well beyond the 1% level; the Fs were, respectively, 26.76 ($df. = 2$ and 141) and 1.84 ($df. = 46$ and 3243).

The 0-delay test of Experiment I did not reveal any sizable or lasting effect for the prior delay-treatment. Groups given delays of 0, 1, and 2 trials had almost equal errors on all four test-trials and, though the group with a 3-trial delay had the largest errors on Trials 1 and 2, there were no differences on Trials 3 and 4.

Trends within- and between-groups were not as clear-cut for algebraic as for absolute error. Algebraic errors gave some evidence of learning, with a grand mean undershoot of about 2° of arc on KR-trial 1 as compared with an overshoot of about 0.7° on KR-trial 3. There was also a slight increase in algebraic error and in trial-to-trial fluctuation for the greater trial-delays. There was no appreciable change in the magnitude of algebraic error after KR-trial 6, however, and, with positive and negative errors cancelling to nearly zero, differences between groups were inconsistent and of trifling magnitude.

Summary. Two experiments are reported which were designed to make delay in knowledge of results (KR) for a simple positioning response analogous to lag in a continuous tracking task by delaying KR over a number of trials. This manipulation of delay, or lag in feed-back, was effective in increasing error (a) relative to a 0-trial delay and (b) progressively as trial-delay increased. The findings point to the importance of intervening responses in determining the effectiveness of delay in KR.

* E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 267-273.

EFFECT OF STIMULUS-DETERMINATION ON RESPONSE-INDEPENDENCE AT THRESHOLD

By JACK E. CONKLIN AND ROBERTA SAMPSON, Montana State College

That human beings are unable to generate a random process has been observed by several investigators. The conclusions of some early experiments with lifted weights, for example, indicate that Os tend to avoid response-repetition,[1] and a tendency to alternate is manifest also in studies of binary guessing responses.[2] Senders has recently reviewed this early literature.[3]

Sequential analysis of threshold-responses has been a subject of further investigation in recent years.[4] Wertheimer, for example, obtained successive threshold-responses for auditory, visual, and algesic stimuli presented successively every 6 sec.[5] He found that threshold-responses were neither independent nor randomly ordered. For longer time-intervals between stimuli (1 min., 3 min., and 1 day), Wertheimer found a cyclical variation in sensitivity. "That the *patterns* of temporal drift of the thresholds of two modalities of an S are similar over the same period of time" also was observed.[6] Collier specifically studied association between successive responses at threshold as a function of the inter-trial interval.[7] Visual stimuli

[1] Leon Arons and F. W. Irwin, Equal weights and psychophysical judgments, *J. exp. Psychol.,* 15, 1932, 733-756; S. W. Fernberger, Interdependence of judgments within the series for the method of constant stimuli, *ibid.,* 3, 1920, 126-151; M. G. Preston, Contrast effects and the psychophysical judgments, this JOURNAL, 48, 1936, 389-402; Contrast effects and the psychometric functions, this JOURNAL, 48, 1936, 625-631; W. D. Turner, Intra-serial effects with lifted weights, this JOURNAL, 43, 1931, 1-25.

[2] A. W. Bendig, The effect of reinforcement on the alternation of guesses, *J. exp. Psychol.,* 41, 1951, 105-107; L. S. Goodfellow, A psychological interpretation of the results of the Zenith radio experiments in telepathy, *ibid.,* 23, 1938, 601-632; M. E. Jarvik, Probability learning and a negative recency effect in the serial anticipation of alternative symbols, *ibid.,* 41, 1951, 291-297; R. L. Solomon, A note on the alternation of guesses, *ibid.,* 39, 1949, 322-326; B. F. Skinner, The processes involved in the repeated guessing of alternatives, *ibid.,* 30, 1942, 495-503.

[3] V. L. Senders and Ann Sowards, Analysis of response sequences in the setting of a psychophysical experiment, this JOURNAL, 65, 1952, 358-374.

[4] G. H. Collier, Intertrial association at the visual threshold as a function of intertrial interval, *J. exp. Psychol.,* 48, 1954, 330-334; J. E. Conklin, Senders on response-sequences, this JOURNAL, 67, 1954, 263-265; J. F. Jerger, On the independence of successive responses in the quantal psychophysical method, this JOURNAL, 68, 1955, 145-147; Senders and Sowards, *op. cit.,* 358-374; Senders, Further analysis of response-sequence in the setting of a psychophysical experiment, this JOURNAL, 66, 1953, 215-228; W. S. Verplanck, G. H. Collier, and J. W. Cotton, Non-independence of successive responses in measurements of the visual threshold, *J. exp. Psychol.,* 44, 1952, 273-282; Verplanck, Cotton, and Collier, Previous training as a determinant of response dependency at the threshold, *ibid.,* 46, 1953, 10-15; Michael Wertheimer, An investigation of the 'randomness' of threshold measurements, *ibid.,* 45, 1953, 294-304.

[5] Wertheimer, *op. cit.,* 294-304.

[6] *Ibid,* 302.

[7] Collier, *op. cit.,* 330-334.

were presented successively at intervals of 1, 2, 3, 4, 5, 12, and 30 sec. Association diminished as the interval increased, rapidly at first, and then more slowly. No cyclical variation in response was reported.

It has been suggested that these experiments invalidate the psychophysical assumption of response-independence at threshold.[8] An alternative interpretation—that, where there is no possibility for stimulus-determination, a *guessing* rather than a *judgmental set* is established—may be made on the assumption that guesses are more likely to be influenced by preceding responses (success and failures) than are sensory judgments. This interpretation, affirmed by Conklin in a preliminary study,[9] has been supported by the data of Jerger.[10] The experiments reported here were conducted to test the hypothesis that response-dependency at threshold is inversely related to the degree of stimulus-determination in the sequence of presented stimuli.

Method. The experimental method required Os to compare five different sets of weights with a standard weight of 100 gm. The variable weights for each set were as follows: Set A, five weights of 100 gm.; Set B, 96, 98, 100, 102, 104 gm.; Set C, 92, 96, 100, 104, 108 gm.; Set D, 84, 92, 100, 108, 116 gm.; Set E, 68, 84, 100, 116, 132 gm. On the assumption that the Weber ratio is approximately 0.02,[11] the differences between weights for Sets A through E are 0, 1, 2, 4, and 8 DL, respectively. According to the hypothesis tested, the effect of preceding responses on threshold-judgments should be maximal for Set A. Responses at threshold should become increasingly random as the discriminability of standard and variable weights increases.

Materials. The weights consisted of small black bottles (7 cm. high and 3 cm. in diameter) packed with lead shot and cotton. The bottles were carefully weighed on a sensitive balance. Additional materials included a metronome, an arm-rest with supporting leather straps, and a blindfold.

Observers. Five undergraduate students (2 men and 3 women) served as Os for Experiment I; 60 students (41 men and 19 women) served in Experiment II. All Os were unfamiliar with psychological experiments in general and the purpose of this study in particular.

Experiment I. Each O was required to make a total of 5000 judgments, 1000 for each set of weights. Two 90-min. sessions per week were given to each set of weights. Each O therefore served in 10 sessions over a period of 5 weeks. The order of presentation of the different sets of weights was determined by a table of random numbers.

During a given experimental session, stimuli were offered in sequences of 25 trials (125 judgments) followed by a 5-min. rest-period. The standard and variable were presented first equally often, but the order was randomized over the eight sequences. An additional restriction consisted of so planning the order of stimulus-presentation (after randomization) that each variable weight was presented an equal number of times as the 1st, 2nd, 3rd, 4th, and 5th weight. Only the effect of preceding judgments on threshold-responses were considered in the final analysis.

[8] Senders, *op. cit.*, 228.
[9] Conklin, *op. cit.*, 263-265.
[10] Jerger, *op. cit.*, 145-157.
[11] J. P. Guilford, *Psychometric Methods*, 1936, 191.

Ordering the presentation of stimuli in this manner allowed an equal N for each O in computing the autocorrelation for four values of *tau*.

Experiment II. The design of Experiment I called for each O to be employed under every experimental condition. In the second study, 60 Os were randomly assigned to three experimental groups, *i.e.* 20 Os to judge Set A, 20 for Set C, and 20 for Set E. Each O was required to make 250 discriminations in two sequences of 25 trials with a 5-min. rest-period between sequences. This procedure yielded a total of 5000 judgments for each set of weights, as in Experiment I. The order of stimulus-presentation was so planned that the threshold-variable would be presented as the 1st, 2nd, 3rd, 4th, and 5th weight an equal number of times. The order of presentation for the weights above and below threshold was randomized.

Procedure. The Os were told that the experiment was designed to test human sensitivity in discriminating weights. Those engaged in Experiment I were informed that they would receive a different set of weights to judge each week. The Os were seated, arm strapped in the arm-rest, and the blindfold secured. They were instructed to judge whether the second weight was 'heavier' or 'lighter' than the first for each pair of weights presented. The time-intervals were governed by a metronome. E so presented the weight that it could be held by the index finger and the thumb for a period of 3 sec. The time-interval between the standard and variable weight was held constant at 1.5 sec. The time-interval between successive pairs was 4.5 sec.

Results. The basic data of Experiment I—frequency of heavier-than-standard responses for each weight-position—were subjected to an analysis of variance. The variances for Os and for weight-position were highly significant ($p < 0.001$). The number of heavier-than-standard responses was not significantly different for the different weight-sets. Two double-interactions—weight-set $\times$ weight position, and Os $\times$ weight-position—were significant ($p < 0.001$). Fig. 1 illustrates the interaction between weight-set and weight-position. It is clear from this graph that every set of weights differed consistantly from every other set in degree of discriminability.

The effect of preceding responses on threshold-judgments was tallied separately for each trial of five judgments. Since all weights were equal to the standard for Set A, the autocorrelations computed for each O are based on an N of 200 for all values of *tau*. N differs, however, for the Sets B through E (160, 120, 80 and 40 for *tau*-values of 1, 2, 3, and 4, respectively).

Group-data are illustrated in Table I. The level of significance is determined by the relationship $X^2 = N\Phi^2$ and the chi-square distribution tables.[12] Only three significant autocorrelations are apparent, Set A at *tau* = 3, Set B at *tau* = 4, and Set E at *tau* = 1. Analysis of autocorrelations for Os separately reveal that the total number of significant autocorrelations for Sets A through E are 6, 3, 1, 2, and 2, respectively. Only one of the five Os shows no significant autocorrelations for any of the weight-sets and *tau*-values.

In Experiment I, it is possible that Os maintain a continuous judgmental set throughout the experiment, since experience with stimulus-determination was available with four of the five weight-sets. This sequential effect between experimental

---

[12] H. M. Walker and Joseph Lev, *Statistical Inference*, 1953, 272.
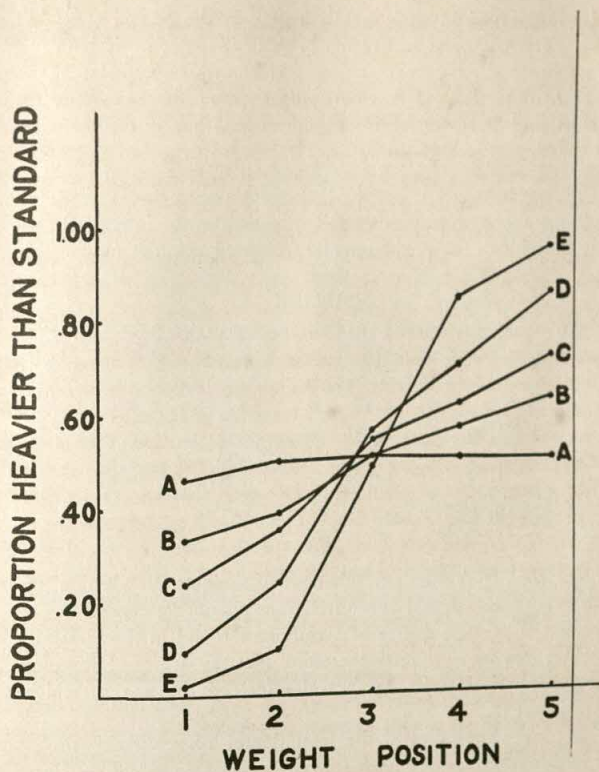
FIG. 1. THE DISCRIMINABILITY OF THE FIVE SETS OF WEIGHTS IN EXPERIMENTAL I

TABLE I

AUTOCORRELATIONS FOR THE COMBINED DATA OF EXPERIMENT I

Set			Tau		
	0	1	2	3	4
A	1.00	0.044	−0.004	0.104*	0.060
B	1.00	0.045	0.007	0.051	0.141†
C	1.00	0.040	0.047	0.081	0.060
D	1.00	0.010	0.027	0.030	0.103
E	1.00	−0.075†	0.007	0.080	0.121

* Significant at the $\frac{1}{10}\%$ level.
† Significant at the 5% level.

conditions (different weight-sets) is eliminated in the second study by assigning different Os to Sets A, C, and E. In spite of this difference in design, the results are comparable to those of Experiment I. Table II shows the autocorrelations obtained for the combined data. For Set A, two significant autocorrelations are indicated for *tau*-values of 3 and 4. One significant autocorrelation is obtained for Set E at *tau* equal 2.

Discussion. The hypothesis tested—that threshold-responses depend upon preceding judgments when stimulus-determination is absent—is partially supported by the fact that individual Os (Experiment I) showed a larger number of significant

TABLE II

Autocorrelations for the Combined Data of Experiment II

Set	Tau				
	0	1	2	3	4
A	1.00	0.056	0.052	0.064†	0.084*
B	1.00	0.050	0.033	0.030	0.060
C	1.00	0.065	0.081†	−0.040	0.060

* Significant at the $\frac{1}{10}\%$ level.
† Significant at the 5% level.

autocorrelations for Set A than any other set of weights. In addition, the combined data in both experiments reveal autocorrelations significant beyond the 1% level for Set A, whereas the greatest degree of significance found in the remaining Sets is at the 5% level. It may be concluded, therefore, that the degree of stimulus-determination in a psychophysical series has a real effect on the way in which threshold-responses are ordered.

One of the Os, who indicated the least degree of sensitivity also showed less randomness in his threshold-responses. He contributed a total of six significant autocorrelations with maximal response-dependence occurring under standard and variable equivalence (Set A). A comparison of the function of autocorrelation of the random signal with simulated records shows a close correspondence for the better trackers and less correspondence for the poor performers. The importance of using trained Os for precise psychophysical determinations (stressed by the early psychophysicians) is borne out by this analysis.

Summary. Two experiments were designed to test the hypothesis that threshold-responses are governed in part by the degree of stimulus-determination in the psychophysical series. It was assumed that dependence of threshold-responses would be maximal under conditions of stimulus-equivalence and would decline as the degree of stimulus-determination increased. The psychophysical series consisted of five sets of variable weights with differing discriminability from the standard. Responses at threshold were analyzed by the autocorrelational method for binary data. The results suggested that two factors are responsible for nonrandomness of threshold-responses: (1) Threshold-responses are nonrandom when stimulus-determination is reduced to a minimum. (2) Untrained Os, or those who show poor discriminative ability, are less likely to generate a random process at threshold.

AVERSIVE BEHAVIOR FOLLOWING LESIONS OF THE SEPTAL REGION OF THE FOREBRAIN IN THE RAT

By W. H. Tracy and J. M. Harrison, Boston University

In two recent papers, Brady and Nauta reported an extensive study of the effects of bilateral lesions of the septal region upon behavior in the rat.[1] Postoperatively, the animals presented a characteristic behavior-syndrome (septal syndrome) which consisted of a number of sensory and motor changes.

About 30 septal animals have been prepared in this laboratory, and their post-operative behavior was similar to that described by Brady and Nauta. The careful examination of these animals, however, revealed differential sensory changes. Squealing and biting could be elicited by the weak stimulation of the back and sides with a blunt probe, but not by the similar stimulation of the ventral surface of the body or the feet. The response obtained from the stimulation of the back was qualitatively similar to responses obtained by strong electrical stimulation of the same area in a normal animal. Squealing was obtained to such weak sounds as general conversation in the vicinity of the animal's cage, while clapping the hands produced the greatly exaggerated startle-reaction described by Brady and Nauta. Vision appeared to be little affected by the lesions. These observations suggested that the operation increased the aversive function of auditory and certain somesthetic stimuli. The purpose of the study reported here was to investigate the differential effects of septal lesions upon the aversive (S^A) and discriminative (S^D) functions of auditory stimuli.

Two experiments were conducted. The purpose of the first was to measure the effects of septal lesions upon both the S^A and S^D functions of auditory stimuli. An animal was able to escape a continually present gray noise by pressing a lever. The relation between the latency of response and the intensity of the noise was measured preoperatively. A postoperative decrease in latency would indicate an increase in the S^A function of the noise. A systematic increase in latentcy could, however, be explained equally well by a decrease either in the S^A or in the S^D function of the noise, i.e. the animal might be deaf. The purpose of the second experiment was to investigate the effect of septal lesions upon the S^D function of the gray noise. Preoperatively, the noise was used as the positive stimulus (S^D) in a discriminative task. If the discrimination was unaffected by the septal lesions, then changes in either direction in the first experiment could be explained by changes only in the S^A function.

EXPERIMENT I

Subjects. Three albino rats, H_{31}, H_{17}, and H_{30}, served as the experimental animals and one, C_{33}, served as the surgical control.

[1] J. V. Brady and W. J. H. Nauta, Subcortical mechanisms in emotional behavior: affective changes following septal forebrain lesions in the albino rat, *J. compar. & physiol. Psychol.*, 46, 1953, 339-346; Subcortical mechanisms in emotional behavior: the duration of affective changes following septal and habenula lesions in the albino rat, *ibid.*, 48, 1955, 412-420.

Apparatus and procedure. Since the apparatus has been described in detail elsewhere, only a brief description is given here.[2] It consisted of a modified Skinner box. Mounted in the top of the box was a speaker (tweeter), through which a hissing (gray) noise of adjustable intensity could be fed. The intensity of the noise was expressed in terms of the peak-to-peak voltage across the terminals of the speaker. The noise was continually present unless terminated for 17 sec. by a lever-press. Responses were registered on a cumulative recorder. Each response also stopped a clock for the 17-sec. silent period. The total time on the clock at the end of a session gave the cumulated latency of all escape-responses.

All animals were initially run at an intermediate intensity of noise until a steady response-strength was attained. At this stage, the control animal was operated. Running of this animal was begun again on the second postoperative day. Following initial training, the experimental animals were run daily at each of three sound-intensities. There was a 1-min. blackout period between the different sound intensities, and each was presented for 15 min. The lowest intensity was approximately the aversive threshold. Running was continued until a minimal day-to-day variation in response-strength was attained, and the animals were then operated. After one to three days, running was begun again and continued for at least sufficient time to establish the postoperative form of the relation between response-strength and the stimulus-intensity.

Surgery and histology. A closed technique developed by Nauta was used in all operations. The cutting instrument consisted of a needle with the point bent at right angles (for a distance of 2.2 mm.) and sharpened on both sides. The instrument was inserted, bilaterally, to a depth of 5 mm. through burr-holes 1-2 mm.

TABLE I

SUMMARY OF OPERATIVE LESIONS

CC, genu of corpus callosum; IG, indusium griseum; LSN, lateral septal nucleus; NS, nucleus of the septum; CF, columns of fornix; PCF, precommissural fornix; FS, fornix superior; FL, fornix longus; VHC, ventral hippocampal commissure; FCH, forward continuation of hippocampus; PT, paraolfactory-tubercular tract. Extent of the lesions: O, less than 10%; 1, 10 to 90%; 2, more than 90%.

Rat	CC	IG	LSN	NS	CF	PCF	FS	FL	VHC	FCH	PT
H₂₅	2	2	2	2	2	2	2	2	2	2	2
H₂₇	2	2	1	2	2	1	2	2	1	1	0
H₂₀	2	2	1	R1, R2	0	0	0	2	0	1	0
H₆₉	2	2	1	2	2	2	2	2	1	1	2
C₂₉	1	1	0	0	0	0	0	0	0	0	0

rostral and 1.5 mm. lateral to the coronal suture. After insertion, the needle was turned to swing the blade medially through a 180°-arc. In the control operation, a plain needle was inserted into the brain in place of the cutting needle. The brains were silver-stained, sectioned, and examined for the locus of primary damage and tissue-obliteration.

Results. The description of the lesion for each animal is given in Table I. The three experimental Ss showed the typical septal syndrome as described by Brady and Nauta. The syndrome was absent in the control S.

---

[2] J. M. Harrison and W. H. Tracy, The use of auditory stimuli to maintain lever pressing behavior, *Science,* 121, 1955, 373-374.

Aversively maintained behavior was at first completely suppressed and then partially suppressed for an extended period. The period of complete suppression was 72, 30, and 54 days, respectively, for rats H_{34}, H_{37}, and H_{39}. The performance of a typical animal, H_{35}, is shown in Fig. 1. Responses returned to preoperative strength by the 100th day in H_{35} and the 69th day in H_{37}. This time was not measured for H_{39}. Aversive behavior was virtually unaffected in the control S. These
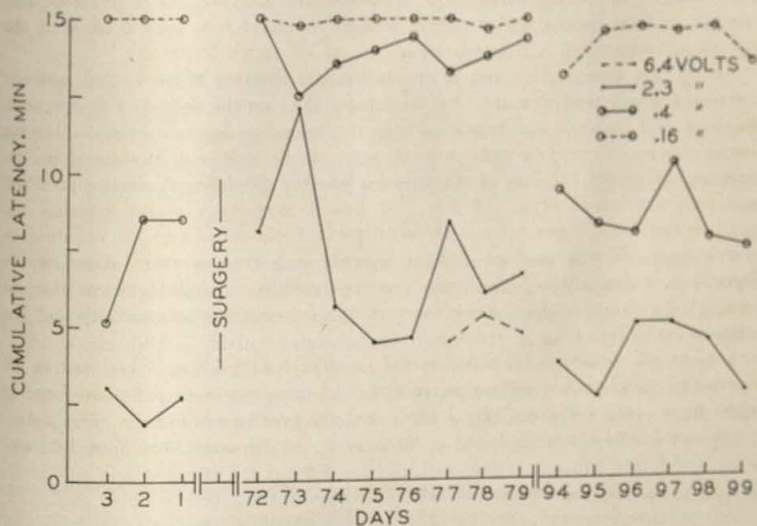
FIG. 1. THE PERFORMANCE OF H_{35} IN EXPERIMENT I

Escape-behavior was completely suppressed until the 72nd postoperative day. The animal was not run between the 79th and 94th days.

results may be explained equally well by the abolition of either the S^D or the S^Δ function of the noise. The second experiment was carried out to determine which of these two functions was most affected.

EXPERIMENT II

Subjects. Two rats, H_{35} and H_{60}, were studied. H_{35} had been used in the first experiment. H_{60} received the same surgical and histological treatment as other experimental Ss.

Apparatus and procedure. A second Skinner box, essentially the same as the first, was used. A discrimination was established between the gray noise of Experiment I (S^D) and silence (S^Δ). The two conditions were alternated at a variable interval of 2.75 min. on the average. Response to S^D was reinforced with food on a 1-min. variable interval schedule, while S^Δ was correlated with extinction.

H_{60} was placed on a 24-hr. feeding schedule and its weight was reduced to 75% of that under *ad lib* eating conditions. The animal was run for 2 hr. a day on the alternating interval and extinction-schedules until a clear discrimination was formed. The intensity of S^D was 0.68 v. for all but the last preoperative session when an

intensity of 2.3 v. was used. The operation was performed and one day allowed for recovery. The animal was run on the second, third, and fourth postoperative days with the same alternating schedules. The intensity of S^D was 0.68 v. for the first two sessions and 2.3 v. for the third. The second animal, H_{25}, was placed on the same feeding and training schedule 42 days after operation. At this time escape-behavior still was suppressed and remained so throughout discriminative training. Daily running on the alternating schedules was continued until a clear discrimination was formed. Intensities of the S^D of 0.68 v. and 2.3 v. were used. Running was discontinued 52 days after operation.

Results. The operative lesions of the Ss are described in Table I. Both showed the typical septal syndrome. H_{69} discriminated clearly on the third and fourth postoperative days and H_{25} readily formed the discrimination and maintained it at intensities of 0.16, 0.4, 0.68 and 2.3 v. Apparently the lesions of the septal region afect neither the S^D function of the stimulus nor the reinforcing property of food.

DISCUSSION

The results of the two experiments, taken together, show that lesions of the septal region markedly diminish the aversive function of auditory stimuli while leaving intact their discriminative function. This conclusion is especially well illustrated by the behavior of H_{25}, which served postoperatively in both experiments. This S *discriminated* the noise during the *same* period in which it was failing to respond to the *identical* stimulus presented in an escape contingency. Postoperatively, none of the experimental animals in the first study gave an extinction-curve. Escape-behavior was completely abolished in the presence of the sound. The immediate implication of this outcome is that the drive-function of the stimulus was abolished. If only the reinforcing function of the termination of the sound had been abolished, a postoperative extinction curve would have been obtained. The initial hypothesis of a postoperative increase in the aversiveness of auditory stimuli has been shown to be incorrect.

The differential sensory effect of the lesions seems to explain the relative lack of postoperative changes in the conditioned emotional response (*CER*) reported by Brady and Nauta in both their studies. In the *CER*, the signal (S^D) for the onset of the electric shock was a clicking noise delivered through an earphone. Postoperatively, septal animals showed either no reduction or only partial reduction in the *CER* (acquired preoperatively) as a function of the locus of the lesion. In other words, responses dependent upon sound as the discriminative stimulus were only little affected by the operation. This is precisely what would be expected from the present findings. Indeed, the work of Brady and Nauta, together with that reported here, suggests that the discriminative function of auditory stimuli is unaffected whether positive or negative reinforcers are used to strengthen the relevant responses. The strength of the *CER* was measured on the third postoperative day, and it is possible that the slight effects reported were due to general immediate postoperative changes.

Three of the Ss in the initial study of Brady and Nauta were first trained in the *CER* postoperatively. The response was acquired at the normal rate. Analogous to this result was the normal formation of the auditory discrimination by H_{25} in Experiment II. The discrimination was first formed postoperatively at a time dur-

ing which the *S* was failing to respond to the same auditory stimulus in the first experiment.

SUMMARY

Two experiments were performed to determine the differential effect of septal lesions on the aversive (escape from sound) and discriminative functions of auditory stimuli. The preoperatively acquired lever-pressing response, which enabled *S* to escape a gray noise, was abolished for from 30 to 71 days postoperatively, and when lever-pressing began again it did so at a level below the preoperative one. Septal lesions failed to abolish or to prevent the formation of an auditory discrimination in which the gray noise used to maintain escape-behavior was the discriminative stimulus. Septal lesions had no effect upon lever-pressing maintained by food. It is concluded that bilateral lesions of the septal region in the rat abolish the aversive function while leaving intact the discriminative function of auditory stimuli. These findings agree with those of Brady and Nauta with respect to the effects of septal lesions on the conditioned emotional response.

SOME OBSERVATIONS ON THE REINFORCEMENT
OF VERBAL OPERANTS

By WILLIAM CODY WILSON, Harvard University, and
WILLIAM S. VERPLANCK, Stanford University

Greenspoon, in 1950, described an experiment in which the rate of saying 'plural nouns' was increased by a procedure that included the use of a verbal reinforcing stimulus, 'mmm-hmmm;' he found that the Ss were unable to report either the reinforcing stimuli or the change in their behavior.[1] That is, Ss were 'unaware' of the experimental manipulation of their behavior. This study reports two brief experiments concerned with similar verbal behavior and similar reinforcing stimuli: we were trying to find out more about some of the conditions under which the effect described by Greenspoon could be obtained, and to clarify the relationships between changes in the Ss' verbal behavior and the Ss' ability to report either the reinforcing stimuli or the change in behavior. In addition, we wanted to experiment with verbal behavior using other than saying-plural-nouns as a response.

EXPERIMENT I

Procedure. The Es, 15 in number (4 women and 11 men), were students in psychology. The Ss, 17 in number (5 women and 12 men), were fellow students. Experimental sessions were, with two unimportant exceptions, conducted in S's living quarters with only E and S present.

The instructions given S were: "This is a study of the vocabulary of college students. Say words. Do not repeat. Do not count. Do not say sentences." Half the Es used as a reinforcing stimulus a casual 'mmm-hmm' or 'good;' the other half wrote down 'significant words' on the data sheet. (The Es were instructed to make an obvious writing movement.) The responses reinforced were 'plural nouns' and 'adverbs.' S kept at the task until he had said 800 words in all; no reinforcement was given for the first 100 words; one response (either 'plural nouns' or 'adverbs') was reinforced during the next 300 words; no reinforcement was given during the next 100 words; the alternate response (either 'adverbs' or 'plural nouns', respectively) was reinforced during the last 300 words. The number of plural nouns, adverbs, and other words was recorded minute by minute by ticking off on data sheets. S was asked at the end of the experimental session to tell all he could about the experiment.

Results. The total rate of saying words did not differ significantly under the various conditions of reinforcement and non-reinforcement (ranked analysis of variance, $p. > 0.50$).[2] The median rate of saying words was 21 per min., with a range across Ss of 8 to 36.

The rate of saying 'plural nouns' increased under the conditions of reinforcement in 13 of 14 Ss. (Two Es, who started reinforcing 'adverbs,' terminated the experiment when their Ss failed to give any. Thus, they did not undertake to reinforce 'plural nouns.') The probability that such a result could be obtained in the absence of a systematic effect of the experimental variable is 0.002 (Wilcoxon's matched pairs test, using signed ranks).[3] The median increase in rate was 95%, ranging from 6% to 2800%. Quite evidently, both kinds of reinforcement, saying 'mmm-hmm' and 'good,' and writing the word down, were effective. In the first case, there was an increase in rate in 7 of 7 Ss; in the second case, in 6 of 7 Ss. When reinforcement was withheld for 'plural nouns' following the period during which they were reinforced, the number decreased in 8 of 11 Ss (the experimental design did not furnish us with this information in all cases); this result could be obtained by chance in the absence of an effect of the experimental variable only once in 10 times. It is probable that the extinction period of 100 words was too short to indicate the full effect of withholding reinforcement.

The rate of saying 'adverbs' increased under conditions of reinforcement in 6 of the 7 Ss who said any adverbs at all; this is statistically significant at $p < 0.025$ (signed ranks test). Nine of the full 16 Ss had an initial rate of zero for saying 'adverbs' hence reinforcement was impossible, and hence this portion of the experiment could not be carried out; the two Es mentioned before stopped at this point and did not carry out subsequent parts of the experiment. The median amount of increase in the rate of saying 'adverbs' under conditions of reinforcement (for those Ss with an initial rate greater than zero) was 500%, with a range from 35% to 1800%. All three Ss from whom reinforcement for saying 'adverbs' was withdrawn showed a decrease in the rate of saying 'adverbs' during this extinction period.

When the Ss were questioned at the end of the experimental session, 12 of 14 mentioned the reinforcing stimuli; 6 of 7 reported the fact in one way or another, that E had said 'mmm-hmm' and 'good' during parts of the experiment ($p < 0.13$, sign test), and 6 of 7 reported that they noticed that E wrote down certain words during the experiment. These Ss would say something like "I noticed that for a while you liked nouns, and then you didn't care what I said," or "Sometimes I'd hit the 'jackpot,' and you wrote it down." One, who showed positive conditioning, thought that it was 'bad' for a word to be written down. Six Ss reported correctly the response-class that had been reinforced; two Ss reported a subclass or a class related to the reinforced class, and four Ss did not report in any way the reinforced class when questioned. One anthropology student, for example, gave the names of tribes, and reported that "peoples were what E liked"; another, although giving plural nouns, reported that "nouns were what E liked."

None of the Ss made any statement to the effect that this was a conditioning experiment. Of those who exhibited skepticism of it as a vocabulary test, some made such statements as, "You were trying to find out how I thought."

Experiment II

Procedure. In the second experiment, the Es were the same students. The Ss were again, with one exception, fellow students; (13 men and 2 women); the one ex-

---

[3] Where N is sufficiently large, and rankings are possible, Wilcoxon's *signed ranks* test is used; in all other cases the sign test is used.

ception was an older physician. The experiments were, with two exception, again conducted in S's living quarters, with E and S alone.

S's instructions were the same as before. 'Mmm-hmm' and 'good,' both spoken casually, were the reinforcing stimuli. Half the Es reinforced words that had anything. to do with 'travel;' the other half reinforced words that had anything to do with 'living-things.' The Ss said 600 words. No reinforcement was given for the first 100 words, reinforcement was given for the selected response-category each time it occurred during the next 300 words, and no reinforcement was given during the last 200 words. The numbers of words dealing with 'travel,' 'living-things,' and all other words were recorded for S. At the end of the experimental session, he was asked to "tell what you can about the experiment."

Results. The total rate of saying words did not differ significantly under conditions of reinforcement and non-reinforcement (ranked analysis of variance, $p >$ 0.75). The median rate of saying words was 19.25 words per min., with a range from 7.6 to 49.

The rate of saying 'travel-words' and 'living-things-words' increased under the appropriate condition of reinforcement in 14 of 16 Ss. The probability that this result could be obtained in the absence of an effect of the experimental variable is less than 0.005 (signed ranks test). The median increase in saying these words was 175%, with a range from 13% to 7600%. Under the control conditions, *i.e.* saying 'travel-words' when 'living-things-words' was being reinforced, and vice versa, the rates did not change significantly. When saying 'travel-words' and 'living-things-words' are considered separately, the results are 7 of 9 Ss for 'travel-words' gave a median increase of 92%, and 7 of 7 Ss for 'living-things-words' gave a median increase of 600%. When reinforcement was withheld (i.e. an extinction condition), the rate of saying these words decreased in 13 of 16 Ss, with 2 Ss continuing at the same rate. This is statistically significant with $p < 0.005$ (signed ranks). The median amount of decrease was 45%, with the range of decrease from 9% to 80%.

At the end of the experimental sessions, 13 of 15 Ss reported in one way or another the reinforcing stimulus. Ten of the 13 Ss who reported the reinforcing stimulus could not identify correctly the response class reinforced, although 5 of these 10 Ss reported a subclass of or a class partially overlapping the one the E was reinforcing. Examples of this are "geographic locations," or "words which have to do with transportation "when 'travel-words' were being reinforced, or "biological terms" when 'living-things-words' were being reinforced.

SUMMARY AND CONCLUSIONS

Despite the absence of instructions, the verbal behavior of most Ss is under the experimental control of an E who systematically reinforces specific classes of verbal response; this occurs whether or not S 'notices' the reinforcing stimuli, although most do so. S will work for an 'mmm-hmm' or 'good,' or a response by E to S's behavior, such as singling out a word for writing-down. S seem to be acting under self-imposed instructions. Thus, S might say "I knew you liked cities, states, and countries, and so I said them," or, "I knew you seemed to want me to say words that had to do with the names of places."

Both the class of responses chosen by Greenspoon and the classes chosen in these experiments prove to behave as responses. Under conditions of reinforcement,

their rate of occurrence increases (provided, of course, that they have an initial rate of occurrence greater than zero). *S* may, however, often respond with a subclass of, or a response class partially overlapping with, the class arbitrarily singled out for reinforcement by *E*. For such *S*s, the response is saying a more narrowly specified class of words.

Some *S*s who were conditioned remained unable to report either the occurrence of reinforcement, or the response-class that they had been giving at an enhanced rate. They had not noticed either the environmental event on which their behavior was conditional, nor that any one response had occurred more frequently than others. Although showing a change in rate, *i.e.* being conditioned, seems a necessary condition for *S*s to report either the reinforcing stimulus or the response-class, the converse is not true. Most interesting, there is no striking difference between the behavior of *S*s who become aware of these variables and the behavior of those who do not.

In summary, (1) the Greenspoon effect is easily reproduced, (2) under our conditions, most *S*s, who show verbal conditioning, are able to report the occurrence of reinforcing stimuli, and of these, most are able to report the response conditioned, and (3) saying 'adverbs' is an example of verbal behavior that has a very low operant level.

REVERSIBLE FIGURES AND EYE-MOVEMENTS

By CHESTER H. PHEIFFER, University of Houston, SPURGEON B. EURE and CHESTER B. HAMILTON, Southern College of Optometry

It was early believed that eye-movements or changes of fixation are responsible for the alternation of figures of reversible perspective.[1] Later, it was concluded that reversal is due to some central factor.[2] Recently, Mull, Ord, and Locke have expressed agreement with the earlier view.[3] Zimmer, cited by Woodworth, found that the eye-movements follow reversal.[4] Sisson concluded that central factors seem to aid the fluctuation. "In fact," he noted, "a . . . case could be made for the theory that central changes of perception induce the peripheral motor inter-fixation of the eyes."[5]

That reversals precede the eye-movements is well supported. Whether a causal relation exists between the reversal and the eye-movement, and whether the reversal gives rise to an eye-movement which has identifiable characteristics, has not been determined. The purpose of the experiments reported here is to answer these questions.

Expriment I. What kind of eye-movement occurs when the staircase-figure is seen to reverse? Each of the five *Os*, college students, was shown the staircase-illusion printed on a card and instructed to look at the figure until he saw it reverse. When he was able to see the reversal, he was placed before an ophthalmograph in normal photographing position. The staircase-card was placed in the card-holder of the instrument. *O* was instructed to "Look at the card and indicate by tapping on your chair when you see the figure reverse." The ophthalmograph was started and, when *O* tapped, *E* covered one reflex-tube to mark the reversal on the film. Each *O* was allowed a minimum of four reversals before the ophthalmograph was stopped.

Analysis of the film revealed two distinctly different types of eye-movement. The first and most frequent type was the erratic wandering found when fixation is not maintained nor a definite observational task set. This type of movement will be referred to as *typical*. The second and far less frequent type was an *atypical* movement (shown in Fig. 1). Of all the movements ($N = 320$), 7.5% were associated with the reversal-marks, and 87% of all movements associated with reversal-marks were atypical; 15% of the reversal-marks had no movement associated with them. The

[1] E. B. Titchener, *Experimental Psychology,* Vol. 1, Part 1, 1906, 152; J. E. Wallin, *Optical Illusions of Reversible Perspectives,* 1905, 1-330; M. F. Washburn, H. Malloy, and A. Naylor, The influence of the size of an outline cube on the fluctuation of its perspective, this JOURNAL, 43, 1931, 484-489.

[2] R. S. Woodworth, *Experimental Psychology,* 1938, 648-649; E. D. Sisson, Eye movements and the Schroeder stair figure, this JOURNAL, 47, 1935, 309-311.

[3] H. K. Mull, Nancy Ord, and Nan Locke, The effect of two brightness factors upon the rate of fluctuation of reversible perspectives, this JOURNAL, 67, 1954, 341-342.

[4] Woodworth, *op. cit.,* 648-649.

[5] Sisson, *op. cit.,* 309-311.

atypical movements associated with the reversal-marks were alternating in character. If the direction of the first part of the first atypical movement was to the right, the first part of the second was to the left, the first part of the third to the right, and so forth.

Experiment II. With a characteristic movement found to be associated with reversals, it was necessary to determine whether the movement causes the reversal or the reversal causes the movement. The procedure was the same as that of Experiment I except that a fixation-point was placed at the center of the staircase-figure,

FIG. 1. ATYPICAL MOVEMENTS
(The break in one line of each record was produced by *E*, who covered one of the reflex-tubes to indicate *O*'s report of a reversal.)

and *O* was instructed, "Look at the dot in the middle of the card at all times and indicate by tapping each time you see a reversal."

Analysis of the film revealed that 77.4% of all movements ($N = 31$) were associated with the reversal-marks and every movement associated with a reversal-mark was atypical. No movement was associated with 7.6% of the reversal marks. Alternation in the direction of the atypical movement continued to be apparent. Since reversals occur without atypical movements but atypical movements do not occur without reversals, it seems reasonable to conclude that the movements do not cause reversals.

Experiment III. To further restrict eye-movement a pinhole aperture, which assured that no movement could occur and the target still be seen, was used. The ophthalmograph was equipped with a septum extending from the nose-position to the center of the card-holder. The reversible staircase was drawn on the right half of the card used, and its left half was blank. When seated at the ophthalmograph, *O* saw the staircase with the right eye only and the blank with the left eye only. The pinhole-tube was placed before the right eye of *O*, and he was instructed to view as much of the staircase as he could and to report when he perceived a reversal. Practice-trials were run to accustom *O* to seeing reversals through the pinholes with both eyes open. After the practice period, *O* was aligned in the ophthalmograph, a compur shutter was placed over the left reflex-tube, and *O* was instructed to press the shutter twice each time he saw a reversal. The first press closed the shutter, and the second opened it. The consensual movements of the left eye were recorded for five *O*s with a minimum of four reversals per *O*.

The pinhole-aperture did not eliminate eye-movements. Analysis of the film revealed that 47.2% of all eye-movements ($N = 72$) were associated with reversal-marks, and that all movements associated with reversal-marks were atypical. No

movement was associated with 14.7% of the reversal-marks. The alternation in direction of the first part of the atypical movements also was noted.

An atypical movement is associated with reversal to such an extent that whenever it is found it indicates that a reversal has occurred. This atypical movement is present when O is viewing the whole figure with no attempt to control eye-movement (87% associated with reported reversals), or while fixating a point (100% associated with reported reversals), or while viewing the figure through a pinhole-tube (100% associated with reported reversals).

Are reversals generally accompanied by atypical movements? In Experiment I 85%, Experiment II 92.4%, and Experiment III 85.3% of the reported reversals were accompanied by atypical movements.

When the Os were allowed to scan the figure at will, 320 movements were recorded, of which only 7.5% were associated with reversals. When the pinhole-tube was used, 72 movements were recorded, of which 47.2% were associated with reversals. When the fixation-point in the center of the field was used, 31 movements were recorded, of which 77.4% were associated with reversals. These data indicate that a fixation-point is more effective than a pinhole in controlling eye-movement and that, when eye-movements are controlled, more atypical movements are associated with reversals and vice versa. Whether the miscellaneous movements interfere with the atypical movement, or whether the control of fixation enables more accurate observation, or both, is at present unknown.

To determine whether reversal preceded or followed atypical movement, the average time from the beginning of a movement to the reversal-mark on the film was calculated. This time was found to be 0.25 sec., which means that the actions required to interrupt the light-reflex must have occurred within 0.25 sec. if the eye-movement preceded the reversal. The time required for O to note the reversal and tap on his chair was not attainable. The time required for E to note the signal and cover the reflex-tube was on the order of 0.75 sec. Consequently, the total time elapsing from the noting of the reversal until it was recorded on the film must have been more than 0.75 sec. It is evident that these processes must have been started before the eye-movement and that the reversal therefore preceded the movement.

Why is this particular movement associated with reversal and how can its characteristics be explained? The most appropriate explanation is that when a figure reverses there is a shift in its apparent position in space. This apparent change in position instigates ocular pursuit-movements characterized by an oscillation of the eyes with decreasing amplitude (Fig. 1) as O becomes adjusted to the new way in which the figure is perceived. A final resolution of the apparent and physical conditions is achieved when O is again fixating that part of the figure determined by his own instructions or by a specific fixation-target. The number of cycles occurring within an atypical movement varies with the individual. Some Os quickly adjust to the new condition with a rapid approximation and correction, while others make a number of corrective movements before steady fixation is regained.

On the basis of this analysis, the question of why there is an alternation in direction of the first part of the atypical movements can be answered. As the staircase reverses, it is seen to move upward and to the right, thus giving rise to ocular movements which are upward and to the right. When the figure returns to

the proper position for a staircase, the apparent shift is downward and to the left. Thus, knowing the way in which the figure is first perceived makes it possible to predict the direction of the first part of the first and of succeeding atypical movements. Consequently, the reversal not only precedes the eye-movement but it also determines the type of movement that will occur.

In supplementary experiments, it was determined that an individual who is familiar with the characteristics of an atypical movement can determine with considerable accuracy from an ophthalmographic record the number of reversals reported by O, and that the atypical movement is associated with the reversal of figures other than the staircase (Tiery's figure, Necker's cube, Mach's book, and the tube composed of circles). Whether an atypical movement is associated with the reversal of figures which are not characterized by an apparent displacement in space (such as the Maltese cross) is not known.

Conclusion. These findings support the view that eye-movement is a result rather than a cause of reversal in reversible figures.

A QUANTITATIVE INDEX OF STIMULUS-SIMILARITY
PROXIMITY VS. DIFFERENCES IN BRIGHTNESS

By Julian Hochberg and Albert Silverstein, Cornell University

The problem of defining or measuring 'similarity' has long been important in perception, in learning, in thinking, and in concept formation.[1] The operational indexing of similarity has been attempted in the fields of human learning and performance, verbal meaning, and discriminative learning of animals.[2] In perception, however, little has been done. A method of indexing is reported here, with a brief consideration of its limitations and difficulties.

A score for stimulus-similarity is obtained along one restricted dimension by setting the 'organizational' effects of similarity into conflict with those of stimulus-proximity; proximity being directly measurable. Consider a matrix of black dots on a white background, with the distances a between rows, b between columns 1 and 2, 3 and 4, 5 and 6, and c between columns 2 and 3, 4 and 5. If $a = b = c$, the dots are equidistant horizontally and vertically. Except for 'space errors' or visual anisotropy, the dots should not be organized into vertical arrays (columns) any more than into horizontal arrays (rows). If we hold a constant and decrease the ratio b/c (i.e. move columns 1 and 2, 3 and 4, 5 and 6 closer together, columns 2 and 3, 4 and 5 further apart), beyond some transitional point, Ss will see the dots arranged in vertical double-column arrays (1 and 2, 3 and 4, 5 and 6), by the operation of the 'law of proximity.' There will in general be a threshold-ratio for this perceptual transition, $b/c = L_o$. Let us assume that the strength of this factor toward vertical organization, for any ratio L_i, is some function, s, of the difference $L_i - L_o$, i.e. $s(L_o - L_i)$.

If we now so change the matrix of dots that the alternate rows, a_1, a_3, a_5, differ along some dimension, relevant to perceptual organization, from the intervening rows, a_2, a_4, a_6, it is no longer the case that the Ss should be as likely to see vertical as horizontal arrays at the previous transitional point: the difference (or departure from similarity) between rows $a_{1, 3, 5}$ and $a_{2, 4, 6}$ should, as it increases past some threshold value D_o, generate horizontal single-row arrays, $a_1 \ldots _6$, by the operation of the 'law of similarity.' Let us assume that the strength of this factor toward horizontal organization, for any stimulus-difference, D_i, is some function k of the magnitude of the difference $D_i - D_o$, i.e. $k(D_i - D_o)$. (It seems likely that a log relationship, following Fechner's *Maasformel*, or Guilford's power law will prove better functions for L, D, or both.)[3]

[1] W. D. Ellis, *A Source Book of Gestalt Psychology*, 1938, 75, 84; E. J. Gibson, Retroactive inhibition as a function of degree of generalization between tasks, *J. exp. Psychol.*, 28, 1941, 93-115; J. S. Mill, *Analysis of the Phenomena of the Human Mind*, 1878, 93-123.

[2] Gibson, *op. cit.*, 93-115; C. E. Osgood, *Method and Therapy in Experimental Psychology*, 1953, 522-537; K. W. Spence, The nature of discrimination learning in animals, *Psychol. Rev.*, 43, 1936, 427-449.

[3] J. P. Guilford, A generalized psychophysical law, *Psychol. Rev.*, 39, 1932, 73-85.

Thus, with $b/c = L_o$, but at some $D_i > D_o$, the Ss should no longer report the transition to vertical organization; instead, the ratio b/c would have to be reduced further, to some value L_i such that $s(L_o - L_i) > k(D_i - D_o)$. As we decrease the similarity between alternate rows, we must therefore increase the factor of proximity in order to shift perceptual organization from horizontal to vertical arrays, and proximity can therefore be used to measure at least certain aspects of similarity.

Procedure. To test this hypothesis, the following apparatus was constructed (Fig. 1): an upright white card (12 in. long × 10⅝ in. high) bore alternate columns 1, 3, and 5, 42 mm. apart horizontally, each composed of squares (6 mm. square) of black (Hering gray paper No. 48) alternating, one beneath the other, 18 mm. apart vertically, with squares of gray (in Condition 1, No. 17; in Condition 2,
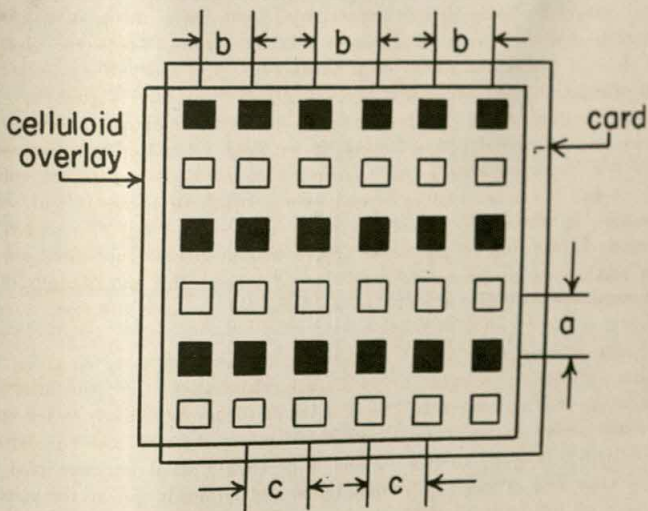
FIG. 1. APPARATUS USED TO VARY THE ARRAYS

No. 6), for a total height of 8 squares in all. Identical intervening columns 2, 4, and 6 are so placed on transparent celluloid overlays atop the card that the ratio b/c, between these columns and those (1, 3, and 5) on the card, can be varied. For each of the two degrees of dissimilarity between dots (Condition 1: black, No. 48 vs. gray, No. 17; Condition 2: black, No. 48 vs. gray, No. 6; conditions presented in balanced order), 11 Ss using a reduction screen found the transitional points by the method of adjustment (10 points in each condition, 5 in each direction) between horizontal and not-horizontal arrays (L_h) and between vertical and not-vertical arrays (L_v). The results appear in Table I.

Results. The results indicate clearly that the use of proximity to score some

aspect of similarity (or dissimilarity), along the dimension of brightness, is possible. Certain limitations should, however, be pointed out:

(1) This technique can index similarity along only those dimensions which will affect perceptual grouping (and thereby discover which stimulus-dimensions are involved in this kind of perceptual organization).

(2) Although there is obviously at least gross ordinal agreement with a *j.n.d.* index of similarity for the values tested (*i.e.* the Hering Gray numbers), we do not know the relationship between the present index and other possible indices.

(3) In some cases, an additive approach to interacting variables, such as simultaneous contrast, will probably not work: for example, if we were to attempt to

TABLE I

Transitional Points for Conditions 1 and 2

(L$_h$ and L$_v$ here represent the mean distance set by each S, in centimeters away from b=c.)

Ss	Condition 1		Condition 2		I_h Differences Cond. 2 −Cond. 1	I_v Cond. 2 −Cond. 1
	L_h	L_v	L_h	L_v		
1	4.2	3.3	4.9	4.2	0.7	0.9
2	8.8	6.2	12.1	6.9	3.3	0.7
3	2.4	0.6	3.7	2.6	1.3	2.0
4	3.2	0.4	4.8	2.4	1.6	2.0
5	3.8	0.0	4.3	1.3	0.5	1.3
6	4.4	4.1	5.2	6.8	0.8	2.7
7	2.8	0.2	8.5	1.4	5.7	1.2
8	1.4	1.3	3.8	1.7	2.4	0.4
9	4.6	−0.4	11.1	1.0	6.5	1.4
10	1.6	0.8	2.5	1.4	0.9	0.6
11	7.1	−0.7	10.0	0.3	2.9	1.0
				M$_{diff.}$	2.42	1.29
				t	3.98	5.86

scale similarity of form by this method, differential 'good continuation' between adjacent shapes, and over-all patterning, would probably prove completely inextricable.

Despite these limitations, the technique may prove a useful one for direct application, as employed here. Perhaps more fruitfully, it can provide a quantitative baseline against which the broadest range of stimuli may be scaled indirectly by providing anchorage points for the more flexible 'esthetic methods' of ranking, *e.g.* triads.

CRITICAL FLICKER FREQUENCY, AGE, AND INTELLIGENCE

By Carney Landis and Violet Hamwi, Columbia University

Several years ago, Tanner found that the critical flicker frequency (*CFF*) correlated between —0.15 and —0.51 with the total score obtained on the *ACE* intelligence test.[1] This *CFF* was obtained by varying the dark-period following each intermittent light-pulse. More recently Colgan obtained from a group of 40 men, aged 65-95 yr., correlations between *CFF* and the total score on the Wechsler-Bellevue (*WB*) intelligence test of 0.46; and between *CFF* and chronological age, —0.32.[2] Application of the statistical method of partial correlation indicated that *CFF* and *WB* correlated 0.36 when age was held constant, while *CFF* and age correlated —0.10 when *WB* was held constant. Colgan concluded that "previous studies relating *CFF* to age must be reëxamined to determine whether intelligence was a controlled variable—if it was not, the conclusions of those studies must be qualified accordingly."[3]

Subsequently, Colgan investigated the relationship between *CFF* and the *ACE* intelligence scores with a group of 90 college students, using equipment similar to that of Tanner.[4] In this study he found no significant correlations, but he did not consider that his results contradicted Tanner's previous findings since he used an intermittent light having a 1:1 light-dark ratio.

There are more than a dozen published investigations of the association between *CFF* and age, but there is little uniformity in the results reported, because of the diversity of methods used, number of persons tested, and range of age examined. About all that may be concluded from the studies is that there is a tendency for *CFF* to be negatively correlated with chronological age. Not enough relevant and comparable information has been included in these publications to ascertain whether this tendency is or is not affected by intelligence.

In different studies done during the past several years we have included measurements of *CFF* in combination with other physiological and psychological indicators. We had not previously examined the possible interrelations between *CFF*, age, and intelligence. In view of Colgan's statements we proceeded to review and analyze the data which we had available.

In Table I are given the correlations between *CFF*, age, and either the total *WB* score or the total verbal *WB* score. In addition, the correlations are given for *CFF* and *age* with *WB* held constant, and for *CFF* and *WB* with age held constant. Column *A* gives Colgan's values obtained from a group of 40 men aged 65–95 yr.[5]

[1] W. P. Tanner, A preliminary investigation of the relationship between visual fusion of intermittent light and intelligence, *Science, 112*, 1950, 201-203.

[2] C. M. Colgan, Critical flicker frequency, age, and intelligence, this JOURNAL, 67, 1954, 711-713.

[3] *Ibid.*, 713.

[4] Colgan, The relation between critical flicker frequency and several psychological variables, Unpublished Ph.D. dissertation, University of Florida, 1954, University Microfilms, Ann Arbor, Mich., No. 9706.

[5] Colgan, *op. cit.*, this JOURNAL, 67, 1954, 711-713.

(Colgan did not state whether he used raw or weighted *WB* scores.) It should be noted that most of his correlations reach the 5% level of significance or beyond, and the effect of holding *WB* constant reduces the correlation between *CFF* and age from −0.32 to −0.10.

Column *B* gives a similar correlational comparison based on preoperative data obtained from 32 mental patients (21 men and 11 women whose ages ranged between 20–61 yr.) who constituted the Columbia-Greystone Group I on which the report of Young was based.[6] The weighted verbal and total *WB* were used. No

TABLE I

PRODUCT-MOMENT CORRELATION COEFFICIENTS BETWEEN CRITICAL FLICKER FREQUENCY (CFF), WECHSLER-BELLEVUE INTELLIGENCE SCORE (WB), AND CHRONOLOGICAL AGE, OBTAINED FROM VARIOUS GROUPS OF Os

Variables correlated	Group				
	A	B	C	D	E
CFF—total verbal WB	0.32*	−0.08	−0.18	0.06	0.08
CFF—total WB	0.46†	−0.07	−0.20	—	—
CFF—age	−0.32*	−0.05	0.20	−0.28	−0.36*
Age—total WB	−0.53†	−0.01	−0.33*	0.18‡	0.18‡
CFF—WB (age constant)	0.36*	−0.03	−0.14	0.01‡	0.02‡
CFF—age (WB constant)	−0.10	−0.05	0.15	−0.27‡	−0.35*‡
Number of Os	40	32	37	33	33

* Significant beyond the 5% level.
† Significant beyond the 1% level.
‡ Correlation based on total verbal WB.

correlation reached a level of statistical significance nor did holding *WB* constant alter the correlation between *CFF* and age.

The data obtained by King and Clausen are given in Column *C*.[7] These data were provided by 37 mental patients (26 men, 11 women; age-range 22–57 yr.) from tests conducted before psychosurgery. These correlations were based on weighted *WB* scores. The correlations are somewhat higher than those of Column *B* but only that between age and total *WB* (0.33) reaches the 5% level of significance. The correlation between *CFF* and *age* is not significant whether or not *WB* is constant. (Note that *CFF* and age correlate 0.20; this is the only positive correlation between *CFF* and age we have ever found either in our own investigations or in the literature.)

Columns *D* and *E* give similar correlational analyses of data from 33 psychiatric patients (18 men, 15 women; age-range 16–45 yr.) studied by Landis and Clausen.[8] The verbal portion of the *WB* test had been given to these patients but the performance portion had not. Column *D* was derived from the *CFF* scores obtained with a Krasno-Ivy flicker photometer having a test-patch luminance of 23 ml.,

[6] K. M. Young, Critical flicker frequency, in F. A. Mettler (ed.), *Selective Partial Ablation of the Frontal Cortex*, 1949, 257-263.

[7] H. E. King and J. Clausen, Psychophysiology in N. D. C. Lewis, Carney Landis, and H. E. King (eds.), *Studies in Topectomy*, 1956.

[8] Landis and Clausen, Changes in sensory and motor performances induced by active psychiatric treatment, *J. Psychol.*, 40, 1955, 275-305.

while Column E came from the CFF scores obtained with a modified stroboscopic lamp-arrangement which gave a test-patch luminance of 3.2 ml. None of the correlations in Column D approaches statistical significance. For Column E, however, CFF and age correlate -0.36, which becomes -0.35 when WB is partialed out. These two correlations are significant at the 5% level.

Colgan gave correlations between CFF and each of the WB subtests, reporting that those with Comprehension, Similarities, Vocabulary, Picture Arrangement, Picture Completion, Block Design, Object Assembly, and Digit Symbol all reached a probability-level of 5% or better.[9] Similar correlational analyses for the groups of data on which Columns B, C, D, and E of Table I were based failed to provide a single significant correlation, all values falling between 0.21 and -0.26.

Data obtained by Landis, Dillon, and Leopold from 91 mental patients on whom CFF, Kent EGY intelligence ratings, and ages also are available.[10] The correlation between CFF and age is -0.41 ($p < 0.01$); CFF and EGY, 0.23 ($p < 0.04$); age and EGY, 0.07; CFF and age, with EGY constant, -0.41 ($p < 0.01$); CFF and EGY with age constant, 0.22 ($p < 0.05$). These values indicate that the negative correlation between CFF and age is not affected by EGY, nor is the relationship between CFF and EGY influenced by age.

Misiak determined the CFF of 319 persons aged 7–91 yr., finding for this large group a correlation of -0.52, but noting "that there were CFFs at age 82 as high as at age 7 and there were CFFs at age 7 as low as at age 80."[11] Various investigators have previously published correlations between CFF and age ranging from 0.00 to -0.74, the values seemingly depending on the age-range and age-composition of the group studied.

Summary. Only Tanner and Colgan have reported significant correlations between CFF and indices of intelligence; no confirmation of their results has been found. There is no obvious reason why intelligence should act as a determinant of CFF. There are enough well-established determinants of CFF to suggest that the correlations reported by Tanner and Colgan are examples of random fluctuation.[12]

---

[9] Colgan, *op. cit.*, this JOURNAL, 67, 1954, 711-713.
[10] Carney Landis, Donald Dillon, and Julius Leopold, Changes in flicker-fusion threshold and in choice reaction-time induced by electroconvulsive therapy, *J. Psychol.*, 41, 1956, 61-80.
[11] Henry R. Misiak, The decrease of critical flicker frequency with age, *Science*, 113, 1951, 551-552.
[12] Landis, Determinants of the critical flicker-fusion threshold, *Physiol. Rev.*, 34, 1954, 259-286.

APPARATUS

AN APPARATUS FOR INVESTIGATING THE METHODS HUMANS USE IN SOLVING COMPLEX PROBLEMS

By Melvin H. Marx, Robert A. Goldbeck, and Benjamin B. Bernstein, University of Missouri

The requirements for complex "human problem-solving tasks" have been recently discussed by Ray.[1] This present report describes an apparatus which meets his two key requirements: multiple scoring, and overtness of behavioral sequences. It has the further advantage of being adaptable to any number of problems where S must formulate hypotheses and gather information about the hypotheses in a step-wise manner before achieving a final solution. Such adaptability makes it possible to design problems which are relatively free of the effects of previous knowledge or variations in experience.

The apparatus was originally designed to simulate the critical aspects of the 'trouble-shooting' process in the APQ-24 radar set.[2] An analysis of this process suggested that there were essentially three phases involved: (1) study of the symptoms that originally indicate a breakdown of some part of the set and give rise to hypotheses about possible sources of trouble; (2) checking of these hypotheses by gathering information about the functioning of components of the set; and (3) replacement of the component finally thought to be faulty on the basis of the original hypotheses and the information gathered in checking the sub-hypotheses. This replacement provides a check of the final hypothesis since, if the hypothesis is correct, the system will function properly after the replacement; if the hypothesis is incorrect, there will still be a malfunction.

The general pattern of apparatus construction to be described obviously was designed to simulate these features; as suggested above, however, the apparatus itself is not limited to such an application.

Fig. 1. Sample System to be Used with Apparatus

The apparatus consists of three main panels, corresponding to the analysis of the trouble-shooting procedure into three parts.

Symptom panel. This panel consists of nine lights and a voltmeter. Any one of these items can arbitrarily be made to function correctly or not, as the experimenter desires; also, the 'rules of the game' can be set up arbitrarily so that malfunction of certain of the symptom-lights or meter indicate malfunctioning of certain components in an overall network. Fig. 1 presents a sample of the kind of network that has been used with the apparatus. The numbers (1–9) across the top of the page correspond to the nine lights, and the 'M' corresponds to the voltmeter. These are the molar indications of performance and represent the outputs of the system. S is told that if the system is working properly, the lights will be on and the meter will give a specified reading. He is also told that if any unit of the system is faulty, the symptom items that normally get a signal from that unit will be off.

Checking panel. This panel consists of 91 switches and a green light. If a part of the system associated with the switch S tries on this panel is working, the green light comes on when the switch is thrown. These switches, then, make it possible to confirm or reject hypotheses about the source of trouble in the system. The rectangles in Fig. 1 correspond to the switches on the checking panel (only 90 are used in the diagram). S is told that each unit in the bottom row of the diagram originates a signal and transmits it to the unit immediately above it, and to any unit indicated by an arrow at the top or output end of that unit. Each of the other units transmits a signal to any unit immediately above it, and to any unit indicated by the output arrow. No unit transmits a signal unless all the units that normally send a signal *to* it are functioning properly. The units in the top row transmit a signal to the symptom item immediately above them in the diagram, causing that symptom item to be on.

S may check for the presence of the signal at the output of any unit by throwing the switch on the checking panel labeled for that unit. If the signal is present the green light will come on; if the signal is absent the green light will not come on.

Replacement panel. This panel consists of 104 switches and a red light. The light remains on as long as there is a trouble in the system. S can 'replace' any unit of the system by throwing the switch associated with that unit on the replacement panel. If S replaces the correct unit, the symptom items on the symptom panel that were off will come back on, and the red trouble light will go off on the replacement panel. S should be able to

make the correct replacement after he has found the unit having no signal output but with all of its input signals present; he finds this unit by checking on the checking panel.

All responses of S are recorded by means of two 20-pen Esterline-Angus recorders. Each switch is so connected to a pair of pens, that the number of the switch thrown and the precise time at which it was thrown can be determined from the Esterline-Angus rolls. Thus the overt pattern of S's activity in solving a problem can be studied; latencies and frequencies of various types of errors as well as correct responses are available.

The equipment, when used with diagrams like that presented here, has good reliability ($r = 0.50$ to 0.99, depending on the type of problem used). It thus fulfills another of Ray's suggested requirements for apparatus to be used for study of problem-solving.[3] Other advantages are that repeated measures of the same type of behavior can be obtained; understanding of the relationships within diagrams is not sufficient for immediate solution of further problems of the same type; the number of responses which can be required for solution of a single problem is large. Thus the overall situation should constitute a sensitive test for picking up small effects of independent variables.

The kind of problem used can be varied to accord with the interests of the experimenter, since the nature of the relationship between the parts of the apparatus and the diagrams used in conjunction with it is an arbitrary one. The problems need not therefore be of the 'parlor trick' variety so commonly found in investigations of problem-solving. The only essential requirement for problems to be used with this apparatus is that some set of alternatives be specified which can be put into correspondence with the switches on the replacement panel, and that some kind of structure be imposed on the rest of the system so that the 'yes-no' kind of information given by the symptom and checking panels is such as to restrict the number of possible correct alternatives on the replacement panel.

A detailed description of the apparatus, including circuitry and a list of parts, is available from the American Documentation Institute.[4] A description of the experimental procedure that has been used with the apparatus is provided in the report of the first investigation completed with it.[5]

[3] Ray, *op. cit.*, 134-149.
[4] Order Document No. 4864 from American Documentation Institute, c/o Library of Congress, Washington 25, D. C.
[5] R. A. Goldbeck, The half-split technique applied to problem-solving tasks of varied complexity, unpublished doctoral dissertation, University of Missouri, 1955.

APPARATUS NOTE

A CONSTANT CURRENT STIMULUS-GENERATOR

For the electrical stimulation of both human and animal Ss it is often important to maintain a constant current in spite of marked variations in the resistance of the skin and the resistance at the point of contact. The apparatus to be described will provide stimuli of uniform intensity in a variety of applications, and is relatively simple to build. It is a modification of a stimulus-generator described by Muenzinger that used the now obsolete type 58 tube and that required batteries.[1]

Operating characteristics. The stimulus-generator provides direct current either for a short pulse, as is used in conditioning, or continuously, as in animal experiments utilizing a grid. The current can be varied between 0.1 to 4.0 ma. Within the range of currents frequently used for human Ss (1 to 4 ma.), for each dial setting

TABLE I

STIMULUS-CURRENT (MA.) WITH VARIOUS RESISTANCES BETWEEN
20,000 AND 1,000,000 OHMS

S's resistance (in ohms)	Dial settings of current control			
	A	B	C	D
20,000	0.25	0.51	1.01	4.20
100,000	0.25	0.50	1.00	4.00
470,000	0.24	0.48	—*	—*
1,000,000	0.24	0.47	—*	—*

* Currents this large are generally not used in experimentation with rats. In experiments with human Ss resistance is nearly always below 100,000.

of the current control the difference between the maximal and minimal current was less than 5% when the resistance load was changed from 20,000 to 100,000 Ω. While within the range of currents frequently used for experiments with rats (0.25 to 0.50 ma.) the variation in current for each dial setting was at no time greater than 6% with differences in resistance load between 100,000 and 1,000,000 Ω (See Table I).

Although in Fig. 1 a spring switch (S3) is indicated for a momentary pulse, a variety of timing circuits may be used instead to obtain uniform stimuli of any given duration. In one application a resistance-capacitance circuit controlling a relay was used in place of this switch to provide repeatable stimuli of short duration for a conditioning experiment.

Construction. The 6AU6 pentode has been selected because in the range below 4 ma. the plate current remains constant with a wide change in plate load resistance. In this application S's resistance becomes the plate load of the tube. Since at lower voltage the change in plate current with change in load resistance would be greater, a plate supply voltage of over 500 v. has been used. This value of the

[1] K. F. Muenzinger and F. C. Walz, An examination of electrical current stabilizing devices for psychological experiments, *J. Gen. Psychol.*, 10, 1934, 477-482.
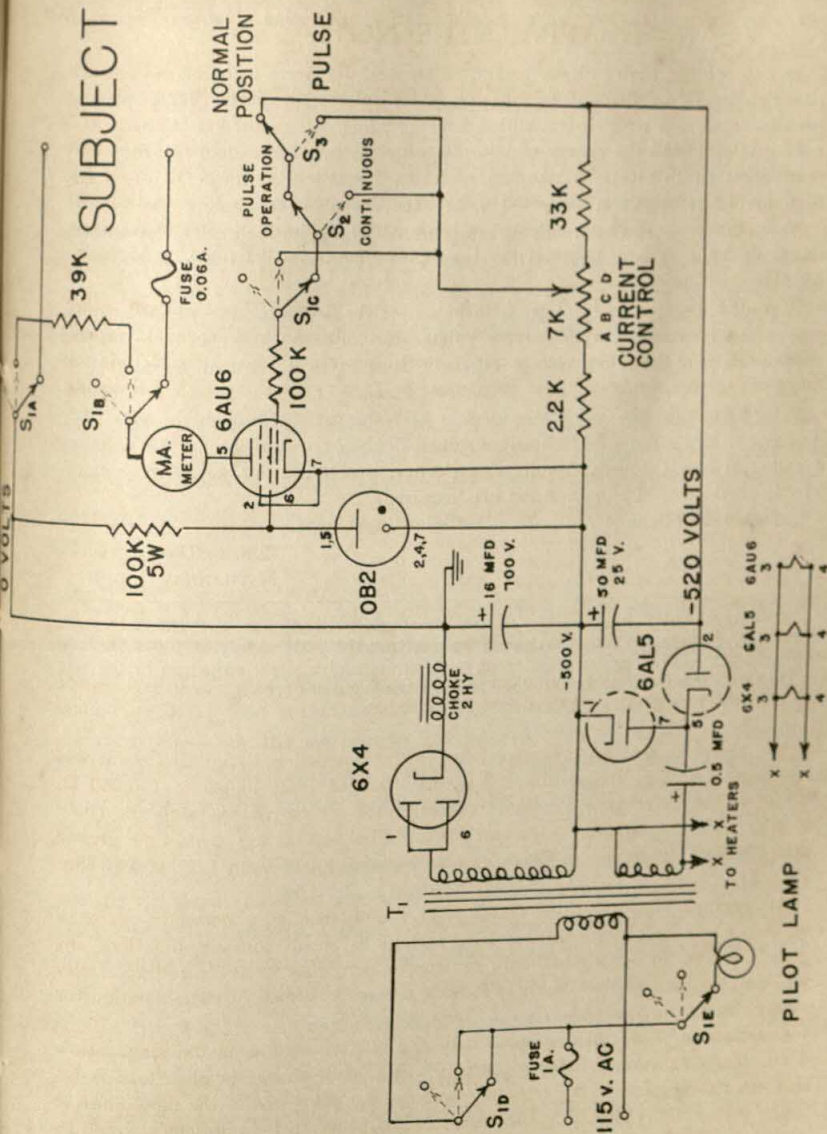
Fig. 1. Wiring Diagram of Constant Current Stimulus-Generator

S_{1A}–S_{1E} are all on common shaft. S_{1A} is shown in operating position; the first position is "off"; the second is for calibration. All resisters are 0.5 w. unless otherwise specified. S_3 is make-before-break momentary contact spring-return switch.

plate supply voltage exceeds 300 v., the manufacturer's rating on this tube. No effects on the tube have, however, been noted and the rating of power dissipation is never exceeded.[2]

Because of the high voltage several safety features were included. The 500-v. point has been grounded and the negative point isolated. In addition the chassis has been mounted in a wooden box with a bakelite panel. All exposed metal parts have been insulated, and the power transformer provides insulation from the line. Care must be taken to choose a transformer whose insulation resistance is sufficiently high to permit operation with 500 v. between the filament winding and ground.

The transformer is a conventional power transformer with high voltage and heater windings. Half-wave rectification has been used to provide the full voltage of the winding.

A regulated screen grid voltage is included for the 6AU6 because without regulation a small variation in the supply voltage can cause a change in the 'constant current' output. An OB2 voltage regulator tube gives a constant screen-grid of 105 v.

A negative grid-bias voltage is used to vary the current through the tube. The bias is obtained from a 6AL5 rectifier-voltage-doubler circuit using the 6.3-v. heater winding as an *AC* source. A voltage of -15.0 v. is obtained this way. A potentiometer is provided to supply the exact bias required.

Bryn Mawr College

ROBERT DAVIDON
NATHANIEL BOONIN

---

[2] If exceptionally high values of S's resistance will not be encountered, more conservative operation can be achieved by reducing the plate supply to 300 v. or less. This change would require a different transformer and a slight reduction in the size of the series resistor connected to the OB2 voltage-regulator tube.

NOTES AND DISCUSSIONS

THE SIZE-DISTANCE PARADOX: A REPLY TO GILINSKY

In the literature of space perception there are three recurrent propositions which are sometimes combined in a theory of size and distance perception.[1] The first proposition is that the perceived distance of an object approaches a limit as its physical distance increases. The second proposition is that the visual angle subtended by an object uniquely determines the ratio of its perceived size to its perceived distance, or

$$\theta = s/d, \dots\dots\dots\dots\dots\dots\dots\dots[1]$$

where θ = visual angle, s = perceived size, and d = perceived distance. This equation has been referred to as the hypothesis of size-distance invariance.[2] Sometimes a variation of the second proposition is combined with a third in the hypothesis that every object has a subjective true size or assumed size, S, which is the outcome of past experience. Then, in place of [1], we write

$$d = S/\theta, \dots\dots\dots\dots\dots\dots\dots\dots[2]$$

in which S rather than s appears. I have previously doubted the value of explaining one subjective magnitude in terms of another, simultaneously perceived, subjective magnitude—sometimes perceived size in terms of perceived distance, and sometimes perceived distance in terms of known or assumed size—but the descriptive accuracy of this approach can be distinguished from its explanatory power. My skepticism first grew out of experimental results which contradicted the second of the three propositions.[3]

Recently, Gilinsky has published a new analysis of those results, purporting to show that they are after all consonant with the three propositions.[4] More specifically, Gilinsky believes that the success of her analysis justifies the particular form in which she has expressed those ideas in her mathematical treatment of space-perception. Of paramount importance in her treatment is the value of the limit, A, which perceived distance is pre-

[1] A. S. Gilinsky, Perceived size and distance in visual distance in visual space, *Psychol. Rev.,* 58, 1951, 460-482.
[2] F. P. Kilpatrick and W. H. Ittelson, The size-distance invariance hypothesis, *ibid.,* 60, 1953, 223-231.
[3] H. E. Gruber, The relation of perceived size to perceived distance, this JOURNAL, 67, 1954, 411-426.
[4] Gilinsky, The relation of perceived size to perceived distance: an analysis of Gruber's data, this JOURNAL, 68, 1955, 476-480.

sumed to approach asymptotically as physical distance increases. On the meaning of A, and on the determination of its value, hinges the whole of Gilinsky's argument. She is clearly aware of this fact, for she has provided seven methods for calculating A.[5] It is the purpose of the present paper to examine Gilinsky's treatment of my results, with special reference to her basic formula for distance and to the calculation of the value of A. The usefulness of this undertaking does not depend merely on the worth of my previously published results, but on the significance of the underlying issues. These issues, too, will be discussed.

Gilinsky's 'basic formula.' The basic formula is $d/D = A/(A + D)$, where d = perceived distance, D = physical distance, and A is the maximal limit of perceived distance for a given O under given conditions. It can be seen that if $d/D = 1$, $D = 0$. Furthermore, if $d/D > 1$, either A or D must be negative, but negative values of A or D are nonsensical. Therefore, it must follow that, under all conditions of judgment, $d/D < 1$. Gilinsky has developed two methods for deriving perceived distance from a set of half-distance judgments, and whenever one method gives $d/D > 1$, the other method gives $d/D < 1$.

Experimental results. Both Gilinsky and I begin our analyses from Equation [1], from which it follows that errors in relative-size judgment should be positively correlated with errors in relative-distance judgment, whether we are comparing individuals or situations. In my experiment, the Os were required to make size-constancy matches when the physical distances of the two objects were in the ratio 1:2. These judgments were coupled with half-distance judgments. For control purposes, the Os also made equal-distance judgments and size-matches when the physical distances of the two objects were equal. The whole procedure was repeated for six distances of the further object.

The experimental results all were contradictory to the hypothesis of size-distance invariance, *i.e.* there was no positive relation between errors in perceived size and errors in perceived distance. (a) There was no correlation between individual errors in size-constancy matches and in half-distance judgments. (b) As physical distance varied, there was no relation between mean errors in size-constancy matches and in half-distance judgments. As the physical distance of the further object increased from 200 to 450 cm., the mean constant error in size-constancy matches rose from 4% to 23%, whereas the mean constant error in the half-distance judgments fluctuated about 17% (Fig. 1). (c) An object which was consistently *under*estimated in relative size was consistently *over*estimated in relative distance.

The third result listed above was labelled the "size-distance paradox" merely because the relation obtained is opposite in sign to that predicted from Equation [1]. For the purpose of comparing two objects, that equation becomes

$$s_1/s_2 = \phi_1 d_1/\phi_2 d_2. \quad\quad\quad\quad\quad\quad\quad\quad\quad [3]$$

Since $s_1/s_2 = 1$, by virtue of the instructions to set the variable object equal in size to the standard,

$$\phi_2/\phi_1 = d_1/d_2. \quad\quad\quad\quad\quad\quad\quad\quad\quad [4]$$

[5] Gilinsky, *op. cit., Psychol. Rev.*, 58, 1951, 460-482.

This is the prediction which follows from Equation [1]. The results show that $o_2/o_1 > \frac{1}{2}$ for the size-constancy matches. This is the usual size-constancy result. When $D_1/D_2 = \frac{1}{2}$, and $o_2/o_1 = \frac{1}{2}$, the objects to be judged are physically equal, but then the further object looks somewhat smaller, and the 0 compensates accordingly in his adjustment of the size of the variable. Meanwhile, the results for the half-distance judgments show that when $D_1/D_2 = \frac{1}{2}$, $d_1/d_2 < \frac{1}{2}$. That is, an object placed at the objective mid-point looks *too near* and 0 compensates accordingly
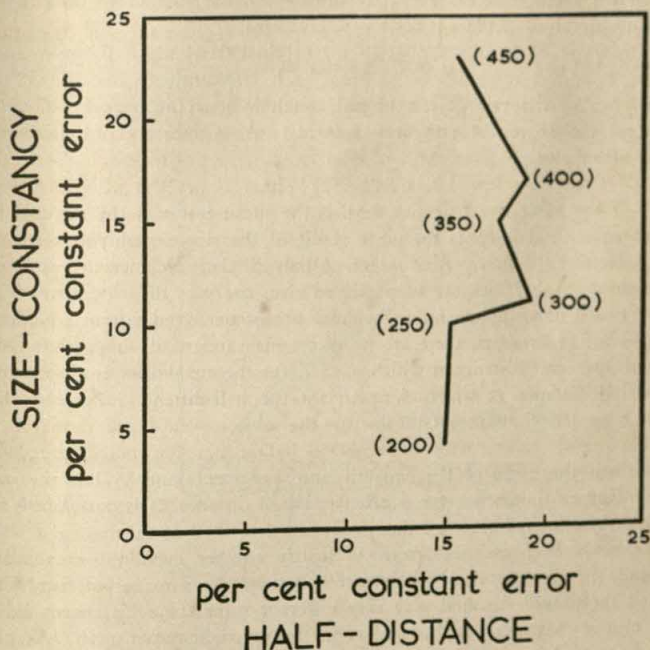
FIG. 1. THE LACK OF RELATION BETWEEN PERCEIVED SIZE AND PERCEIVED DISTANCE

The figures in parentheses are the physical distances bisected in the half-distance judgments and the physical distances of the further object in the corresponding size-constancy judgments.

in his adjustment of the variable distance. But if $o_2/o_1 > \frac{1}{2}$ and $d_1/d_2 < \frac{1}{2}$, then Equation [4] does not hold. Since Equation [4] is derived directly from Equation [1], these data constitute strong evidence against the hypothesis of size-distance invariance. That is the size-distance paradox.

Gilinsky's analysis of my data does not concern itself with this paradox; it deals only with the result (b) listed above—the lack of a positive correlation between errors in size-constancy matches and in half-distance judgments as physical distance increased. The difficulty of the problem which Gilinsky has set herself can be seen from Fig. 1, which sets forth graphically the apparent lack of relation between

the two kinds of judgment. It will be instructive first to examine Gilinsky's own application to my results of the method she set forth in 1951.

Gilinsky's method. Gilinsky chose to calculate the value of A by means of her seventh method, that is "from coupled observations of perceived size and perceived distance." The method is to plot a subjective size ratio s/S against d, and to so extend the graph linearly that it cuts both axes. The x-intercept gives the value of A directly, and the y-intercept gives the value of B/A, a related constant.[6] Although the method seems very straightforward, the step of plotting perceived size against perceived distance is beset with problems.

(1) Definition of perceived size. In my original data, which Gilinsky analyzed, there are no direct measures of s. From each size-constancy match, we get the *physical* sizes of the two objects judged equal. We get no measure of perceived size except the subjective size-ratio, $s_1/s_2 = 1$. As a measure of perceived size, Gilinsky uses the value

$$s/S = ESE/\tfrac{1}{2}ES \dots\dots\dots\dots\dots\dots\dots\dots\dots[5]$$

where S is the subjective true size, ESE is the mean result of the size-match with distances equal, and $\tfrac{1}{2}ES$ is the mean result of the size-constancy matches.

(2) Definition of perceived distance. Although Gilinsky uses the *ratio* of two size-judgments as her measure of perceived size, she uses the *arithmetic difference* between two distance-judgments as her measure of perceived distance. No explanation is given. In my data, there are no direct measures of d. From each judgment we get the *physical* distance at which $d_1 = d_2$ for the equal-distance judgments, and the physical distance at which $d_2 = 2d_1$ for the half-distance judgments. For the measure of perceived distance, Gilinsky uses the value

$$d = ED - \tfrac{1}{2}D \dots\dots\dots\dots\dots\dots\dots\dots\dots[6]$$

where ED is the mean of the equal-distance judgments and $\tfrac{1}{2}D$ is the mean of the half-distance judgments, for a given standard distance. It is particularly strange that Gilinsky should have used this unexplained method of calculating d when she has previously been at great pains to justify another method for calculating d graphically by plotting the half-distance function.[7] As a matter of fact, Equation [6] and the graphic method give very different values of d, as shown in Fig. 2. Using Gilinsky's graphic method, all of my results yield values of $d/D > 1$, which is not permitted in Gilinsky's theory. Using the method of Equation [6], all of my results conform nicely with Gilinsky's theory.[8] Gilinsky cannot have it both ways, for whenever one method gives $d/D > 1$, the other method will give $d/D < 1$. Upon which method are we to rely?

(3) Predicting with A. According to Gilinsky, the value of A may be applied in "predicting" any of the results which were not used in deriving A. Since, in Gilinsky's treatment, all of the data obtained from my Os were used in deriving A, not a very great deal remained to be predicted. Gilinsky asserts, however, "The results thus far analyzed are independent of physical distance, that is to say the physical distances used in the experiment did not enter the calculations on which the numerical values of A and B are based."[9] It should be possible, therefore, to

[6] *Ibid.*, 481-482.
[7] *Ibid.*, 473-475.
[8] Gilinsky. *op. cit.*, this JOURNAL, 68, 1955, 476-480.
[9] *Ibid.*, 478.

predict the relation between perceived size and physical distance by means of Gilinsky's basic equation for perceived size, and, similarly, the relation between perceived distance and physical distance. Gilinsky proceeds to show that her theoretical values and my obtained values are in good agreement. The argument that this really is prediction hangs by a very slender thread indeed. Gilinsky has failed to notice that the obtained values for the mean equal-distance judgments are within 1% of the physical distance being judged. That is, the values of *ED* used in calculating *A* are almost precisely the same as the values of *D* which she is supposedly
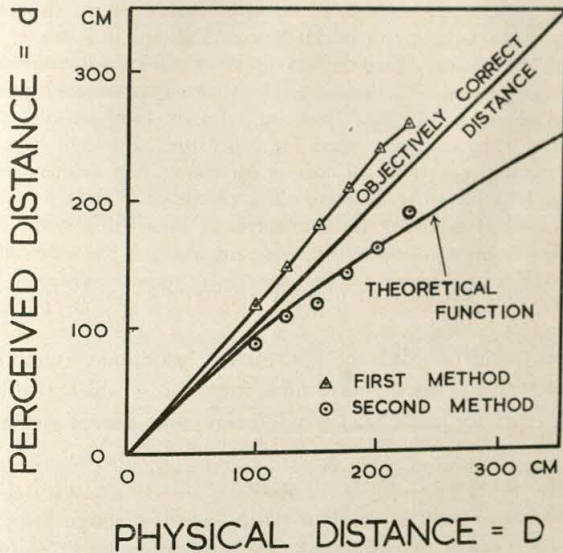
FIG. 2. GILINSKY'S THEORETICAL FUNCTION AND HER TWO METHODS OF CALCULATING SUBJECTIVE DISTANCE

The lower solid line is Gilinsky's theoretical curve, showing a close relation with Gruber's results plotted according to Gilinsky's second method; all values of $d/D < 1$. The upper solid line shows Gruber's results plotted according to Gilinsky's first method; all values of $d/D > 1$.

predicting. It may be that Gilinsky's application of Method 7 merely succeeds in recovering some of the information it utilizes.

The choice of Method 7, beset with the difficulties outlined above, is especially difficult to understand when one learns that Gilinsky has previously developed a method of deriving *A* which is entirely free of these problems.

Method 5. In Method 5,

$$A = D_1 D_2 / (D_2 - 2D_1) \quad \dots\dots\dots\dots\dots\dots\dots\dots [7]$$

where D_2 is a measured distance and D_1 is the measured setting which 0 chooses as the corresponding subjective half-distance.[10] Now some very simple algebra shows

_____

[10] Gilinsky, *op. cit., Psychol. Rev.,* 58, 1951, 481.

that A must be negative if $2D_1 > D_2$. What is the meaning of negative values of A, where A has been defined as the maximal limit of perceived distance? This is not a trivial question: On the one hand Equation [7] is derived directly from Gilinsky's basic formula for perceived distance, and therefore the whole theory is at stake, not merely one method of deriving A. On the other hand, my experimental results for the half-distance judgments are of just this kind, all leading to negative values of A. In addition, where Gilinsky has performed a similar experiment, half her Os $(N = 2)$ gave half-distance judgments similar to those of my Os over the comparable distances.

In short, it appears that Gilinsky's method of analysis is not so clear and consistent as its mathematical form might suggest. Different methods of calculating the fundamental constant A lead to very different results. All except one of the seven methods suggested for calculating A depend on the measurement of subjective size and subjective distance. These measures, in turn, are difficult to make unambiguously. There is one method for calculating A which does not utilize measures of subjective size or distance—in Method 5, only the physical measures corresponding to a subjective distance-ratio appear—but it leads to a nonsensical, negative value of A whenever the constant error in a half-distance judgment is positive in sign.

General issues. Let us turn now to the more general issues which are at stake:

(1) Does perceived distance approach a maximum asymptotically as physical distance increases, following a function in which at all values of D, $d/D < 1$? If so, half-distance judgments must always produce a negative constant error, but in my experiments, employing distances up to 450 cm., the constant error is uniformly positive. In a more recent study, Purdy and Gibson found positive constant errors in half-distance judgments over a much wider range of distances.[11] In another study, Gibson, Bergman and Purdy found the relation between true distance and estimated distance to be linear over very wide ranges.[12] This, too, is contrary to Gilinsky's hypothetical function.

(2) Are errors in perceived size positively correlated with errors in perceived distance? My research produced three major results, each of which argued against such a correlation; at best Gilinsky has challenged only one of these results. In addition, Kilpatrick and Ittelson have reported an ingenious experiment which controverts the hypothesis of size-distance invariance.[13] By passing a small object through a slanted window

[11] Jean Purdy and E. J. Gibson, Distance judgment by the method of fractionation, *J. exp. Psychol.*, 50, 1955, 374-380.

[12] E. J. Gibson, Richard Bergman, and Jean Purdy, The effect of prior training with a scale of distance on absolute and relative judgments of distance over ground, *ibid.*, 50, 1955, 97-105.

[13] *Op. cit.*, 223-231.

seen in reversed perspective, they produced an illusory S-shaped movement of the object in the horizontal plane—although its actual distance from O remained constant, it gave the appearance of varying in distance. The Os were questioned as to any changes in perceived size, and, in approximately half the cases, the results contradicted size-distance invariance. The authors correctly conclude that, since the hypothesis predicts that every change in perceived distance will be accompanied by a corresponding change in perceived size, the hypothesis is incorrect.

(3) Does perceived size decrease in negatively accelerated fashion as physical distance increases? This is really a composite prediction which follows from the joint operation of Equation [1] and the assumption of a maximal limit which perceived distance approaches asymptotically. Oddly enough, over short distances, such as I used, and Holaday used, perceived size did decrease as distance increased, but over the much longer distances of Gibson's experiment there was no such effect.[14] Moreover, as Smith has pointed out, the experiment of Holway and Boring, which Gilinsky employs in her argument, actually showed a *linear* relation between perceived size and physical distance.[15]

Perhaps the hypothesis which has been found wanting should be re-christened the hypothesis of perceived-size-perceived-distance invariance. In spite of the evidence against this hypothesis, it is clear that perceived size depends on some function of physical distance. This function is not given in the law of the visual angle, nor is it given in the determinants of perceived distance. The information we lack can only be supplied by concrete experimental analysis of the *proximal stimulus-correlates of physical distance*. There are probably several sets of such stimuli, some relevant for perceived distance, some for perceived size. The conditions under which these sets may sometimes overlap or coincide suggest a problem for future research, not a foregone conclusion.

Summary. Gilinsky's methods of analysis are shown to be ambiguous and inconsistent. Her mathematical theory of space perception is derived from three fundamental propositions, two of which are shown to be contradicted by a variety of experimental results. In the experiments discussed,

[14] J. J. Gibson, *The Perception of the Visual World*, 1950, 183-186. Gruber, *op. cit.*, 411-426; B. E. Holaday, Die Grössenkonstanz der Sehdinge bei Variation der inneren und äusseren Wahrnehmungsbedingungen, *Arch. ges. Psychol.*, 88, 1933, 419-486.
[15] A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21-37; W. M. Smith, Gilinsky's theory of visual size and distance, *Psychol. Rev.*, 59, 1952, 239-243.

perceived distance does not approach a limit as physical distance increases, nor is physical distance always underestimated. Errors in size-judgment are not positively correlated with errors in distance-judgment.

University of Colorado

HOWARD E. GRUBER

TECHNICAL TERMS IN SCIENCE AND TECHNOLOGY

An article in a recent number of this JOURNAL on "Parallax and perspective during aircraft landings,"[1] puts forward the beginnings of a theory which has been carried much further in the United Kingdom—but by engineers and pilots, not by psychologists. A system of visual aids largely based on this theory has been in operation at a number of airports in Europe since 1948, and is now standard on most instrument-runways there. On these runways, accidents due to visual misjudgments during the final approach and landing have been reduced almost to zero. In contrast to this there are about a dozen different systems presently in operation in the United States, only one of which has a performance comparable with that obtained from the system used in Europe. In order that the standardization so desirable for safety should be achieved, it is necessary that some theory of how visual judgments are made in motion should be generally accepted. If the psychologists are going to join in this work, and it is highly desirable that they should, then I suggest that they begin by adopting the terminology which the engineers and pilots already use. Air transport is a world wide affair, and if psychologists are to influence people in executive positions, their papers must be clear and precise and easily translated into other languages. This is possible only if words are used in their dictionary meanings, or else very carefully defined, but the psychologists working in the field of visual perception seem to be so lax in this respect that their writings have little meaning for the pilot or engineer, and may therefore tend to be ignored. Let me illustrate this by examining the meaning of four words, two of which appear in the title of the above paper. Anyone who thinks that the meanings I give for these are wrong should consult any good technical dictionary.

Parallax. Parallax is a term which the psychologists have borrowed from astronomy. Celestial bodies have in practice to be observed from a point

---

[1] J. J. Gibson, Paul Alum, and Frank Rosenblatt, *op. cit.*, this JOURNAL, 68, 1955, 372-385.

on the surface of the earth, from which it follows that observations made at different observatories at any given moment will differ. To put these observations on a common basis, astronomers are in the habit of referring them to a standard reference-point. Parallax is the angle between the line joining the body to the actual point of observation and the line joining the body to the reference-point. Two reference-points are used, the center of the sphere which contains the equator, and the center of the sun.

When the observer is on the surface of the earth and the reference-point is the center of the sphere which contains the equator, then the parallax varies with the time of day, and has a maximal value known as the 'equatorial horizontal parallax.' If the word is used without qualification, as in the title of the paper referred to above, this is what is meant or should be meant. It is used for bodies which are relatively near.

When the reference-point is the center of the sun, then the parallax varies with the time of year, and has a maximal value known as the 'annual parallax.' It is used for bodies which are relatively distant, *i.e.* stars, and is expressed in 'parsecs.' A star which has an annual parallax of one parsec is at a distance from the sun such that the radius of the earth's orbit subtends one second of arc at the star.

The important point for borrowers of words to notice is that the terms equatorial horizontal parallax, and annual parallax, both refer to the angle between the directions of the body when observed from two points, the line joining which is *at right angles* to the line joining one of these points to the body. The length of this transverse line is the radius of the earth in the first case, and the radius of the earth's orbit in the second case. These terms therefore refer to a change in the apparent direction of a body due to a *transverse* shift of a certain amount by the observer. Opticians, notably Helmholtz, have realized that a movement of the head transverse to the direction in which the observer is looking provides a cue for the distances of objects in the field of view, and many people in their school days have used this effect to find the position of the virtual image formed by some optical system. It is because the word parallax is associated with a transverse movement of a given amount that we in the United Kingdom have been careful *not* to use the term 'monocular motion parallax' to describe the *continuous* changes which occur in the perspective image when the observer is in continuous motion, and is looking in the direction in which he is going. The terms which we use for this are streamer-pattern, instantaneous origin of streamer-pattern (now usually called 'Point X') streamer-velocity, vertical component of streamer-velocity,

(a very important one for landing) and horizontal component of streamer-velocity. If we are referring to changes in the retinal image instead of the perspective image, these terms are preceded by the word 'retinal.' These terms are vivid, almost self-explanatory, and are not vitiated by associations carried over from another science. To use the term 'binocular parallax' to describe the disparity between the retinal images in the two eyes seems to me to make confusion worse confounded.

Perspective. Perspective is the technique of so representing solid objects on a flat plane that when an observer looks at the picture on this plane from a particular point, the retinal image is identical with what it would be if the observer looked at the actual objects from the same point. The investigation of visual problems by the use of such pictures is what we call 'perspective analysis,' and for this purpose we normally use 'parallel perspective.' When any adjective, such as 'linear' or 'parallel,' is used to qualify the word 'perspective,' the resultant term should denote some special variant of the technique, usually a simplified variant. In view of this I regard the term, 'motion perspective,' proposed by the authors of the above paper, as a grammatical curiosity. Incidentally, a perspective picture showing the streamers is far more informative than the diagrams shown in the paper. These merely tabulate angular velocities, and do not reveal what the function of the streamer-pattern really is in the control of a vehicle.

Orientation. 'Orientation' originally meant the relation of a building (usually sacred) to the east, but it now means the relation of any object to the points of the compass. Runways have orientations which are denoted by numbers, *i.e.* a runway with an orientation of 270 runs from east to west. Psychologists in the United States commonly use the term 'orientation in space' to denote the pilot's impression of his lateral distance from the desired path, his height, and the attitude[2] of the aircraft, etc. The term which pilots use for obtaining a combined impression of height and attitude is 'establishing the ground plane,' *i.e.* establishing the relationship of the plane of the ground to the framework of the aircraft.

Texture. When we in the United Kingdom started work on visual aids in 1945, we realised almost immediately that establishing the ground-plane quickly and safely (which is vital for the pilot after an approach on

[2] "Attitude" is the technical term for the angular relationship of the framework of the aircraft to the horizontal plane. "Ground plane" includes height as well as attitude.

instruments), depended amongst other things on the texture of the surface. We therefore defined 'texture' for our purposes as the presence on the surface of solid objects or markings having appreciable dimensions in a plane transverse to the direction in which the observer is looking, and we used the term 'uniform texture' to describe the case in which an object or surface marking with a constant dimension in this plane is repeated at regular intervals longitudinally. This simple conception has been compromised by the use in the United States of the term 'texture gradient.' Gramatically this should mean texture graded in size, as at Chesil Beach in the south of England, where the pebble size gets progressively larger from one end of the beach to the other. As used by Gibson, I think it refers to the appearance of the perspective image of a surface on which the texture is uniform, *i.e.* there is a confusion between the actual surface and its image.

In conclusion there is one further observation which I wish to make on this subject. Everyone who has worked on the man-machine relationship knows that stability in control is very difficult to achieve without using information as to the rate of change of the quantity controlled. No one, however, except a few workers in the United Kingdom seems to have realised that the capacity to control a vehicle in a stable manner, *i.e.* to make it track along a given path without wandering (a thing that most of us can do when driving a surface vehicle), implies that the driver is using rate-information. This is what the pilot of an aircraft is trying to do, but his trouble is that he has to do it in two planes at once with a vehicle which has six degrees of freedom. The streamer-theory of visual judgments in motion *begins* with the idea of rate, and proceeds to show how this vital information (for both planes) is extracted from the visual field. As mentioned above, the application of this theory greatly reduced the accident rate. In addition, it reduced the failures in first attempts to land from more than 30% to less than 2% in a visibility of half a mile and also lowered the operating limit to 500 yd., both of which results are important from the point of view of the traffic controller. The practical dividends to be obtained from this work are therefore very great. If psychologists wish to make a useful contribution, they would be well advised to make themselves familiar with this way of thinking about the visual control of aircraft, and to adopt in the future a less muddled terminology.

Royal Aircraft Establishment   E. S. CALVERT
Farnborough, England

FIGURAL AFTER-EFFECT, AFTER-IMAGE, AND PHYSIOLOGICAL NYSTAGMUS

Osgood and Heyer recently proposed a statistical explanation of figural after-effects (*FAE*) deriving from the Marshall-Talbot theory of *acuity*, which is in good accord with known neurological structure.[1] While radically different from the Köhler-Wallach direct-current explanation, it seems to deal with most of the existing data quite as adequately.[2] Further investigation of the detailed conditions for *FAE* therefore appears to be necessary.

The new theory proposes that a single stimulating contour results in a whole region of excitation due to the two mechanisms of 'neural overlap' and 'physiological nystagmus,' the perceived contour occurring at the peak of excitation in this region. It is with the role of physiological nystagmus that the present paper is concerned. Nystagmus produces a fluctuating locus of the stimulating contour on both the retina and the cortex, resulting in a roughly bell-shaped distribution of excitation over a given time period. In consequence, a band of differentially adapted cells remains after the stimulus is removed. A new stimulating contour, at some distance from the first, will produce a second band of excitation, but insofar as this region overlaps with the band of *adaptation* of the previous stimulus, the peak of excitation in the second region will be shifted away. Thus, perceived stimulus-displacements will result, roughly of a contour-repulsion nature, *i.e. FAE*. The purpose of this experiment is to determine whether physiological nystagmus is necessary for *FAE*.

Method. To obtain *FAE*, there must be an inspection-figure (*I*-figure) which is presented for some period of time, and a test-figure (*T*-figure) which is presented long enough for a judgment. To eliminate the effects of physiological nystagmus, however, a figure ordinarily can be presented only for a fraction of a second. This difficulty can be circumvented by the use of some very involved apparatus, but there is a simpler way. A brightness after-image can be occasioned by very brief exposures, yet maintained, after the exposure has ceased, for relatively long periods, safely ensconced on the retina, immune to physiological nystagmus and other eye-movements.

---

[1] C. E. Osgood and A. W. Heyer, A new interpretation of figural after-effects, *Psychol. Rev.,* 59, 1951, 98-118; W. H. Marshall and S. A. Talbot, Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity, in Heinrich Klüver (ed.), *Visual Mechanisms,* 1942, 117-164.

[2] Wolfgang Köhler and Hans Wallach, Figural after effects, *Proc. Amer. Phil. Soc.,* 88, 1944, 269-357; see, however, John Krauskopf, The magnitude of figural after-effects as a function of the duration of the test-period, this JOURNAL, 67, 1954, 684-690.

Exposure of the *T*-figure (to *O*'s left eye) and *I*-figure (to *O*'s right eye) was provided by photographic flash-bulbs behind cut-out windows; exposure was something less than 1/70 sec. in duration. The *I*- and *T*-figures are shown in Fig. 1. To prevent after-images from fading out too quickly (do they do so normally because of 'self-satiation'?), a slow episcotister intermittently illuminated a viewing screen upon which *O* projected the after-image.

The *O*s of the experimental group first judged which circle of the dimly lit *T*-figure was larger (only one judged the right larger and, for him, the *T*-figure was reversed). They next fixated the dimly-lit *I*-figure, and a

4°	4°
2°5' 1°20'	1°45' 1°40'
I-figure	*T*-figure
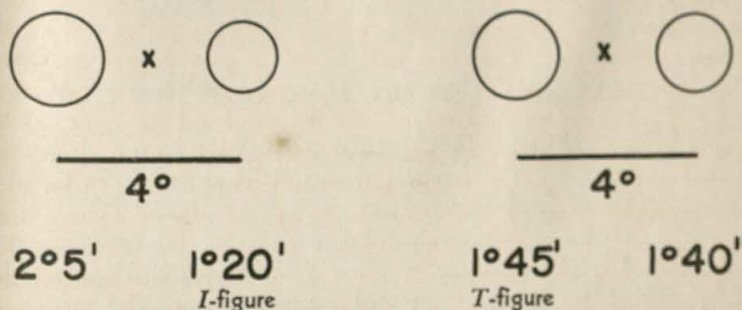
FIG. 1. INSPECTION- AND TEST-FIGURES

photoflash established the *I*-figure after-image, which was projected on the viewing screen for 1 min. The *T*-figure after-images were then established in the same fashion, and projected on the viewing screen. The *O*s were asked which of the two circles of the *T*-figure after-image appeared larger, whether they could still see the older after-images, and whether the old circles were concentric with the new ones. This last point was a check on whether comparable fixation had been achieved for the flash exposures to *I*- and *T*-figures; 4 of the 20 *O*s had to be rejected as failing to meet the fixation-requirements.

Seven *O*s served as a control group to determine whether there might not be some anomalous size-contrast between the after-images of *T*- and *I*-figures, despite their heterocular separation, such simultaneous contrast perhaps mimicking the 'successive contrast' of the *FAE*. The procedure differed from that for the experimental group only in that no inspection-period followed the presentation of the *I*-figure.

Of 16 experimental *O*s retained, 12 judged the right *T*-circle larger, 2 judged the left larger, and 2 judged them equal. In terms of figural after-

effects we should expect the right T-circle to be judged larger after satiation. (It should be noted that the actual sizes of the T-circles were such as to work against this outcome.) All 7 control Os judged the objectively larger T-circle as the larger.

The results indicate that physiological nystagmus is probably not *necessary* for figural after-effects; we cannot yet say whether it is sufficient, and the statistical theory still has another mechanism—the neural spread—which remains untested. The results are neutral with respect to the Köhler-Wallach theory.

<div align="right">

Cornell University   JULIAN HOCHBERG
JOHN HAY

</div>

INTELLIGENCE AND REVERSALS OF PERSPECTIVE

Mull, Arp, and Carlin[1] have reported evidence for the role of 'central' factors as determinants of reversals of perspective as well as for Köhler's isomorphic theory.[2] This evidence took the following forms: (a) although it was found that wide individual differences existed in average rate of uncontrolled cube fluctuations, a given S maintained strikingly similar rates whether the two eyes were used together or separately; and (b) a relationship was reported between intelligence and ability to control reversals.

This note reports the results of an attempt to replicate the relationship between controlled reversals and intelligence found by Mull, et al. It would be of considerable theoretical importance, if a reliable relationship between intelligence and a relatively simple perceptual task, like the ability to control reversals of perspective, could be established. In view of the small sample and the borderline significance of the correlation between Necker cube control scores and the Scholastic Aptitude Test ($r = 0.33$, $df. = 38$, $p < 0.05$) in Mull's study, further confirmation was considered necessary.

The procedure of Mull, Arp, and Carlin was followed closely. Forty-three men, undergraduates in an introductory course in psychology, were individually presented at a distance of 1 m. with a drawing of a Necker cube (photographically reduced to 5 cm.). S sat with his head in a chin-and-head-rest. After a brief practice trial, two trials of 2 min. each (separated

[1] H. K. Mull, K. Arp, and P. Carlin, Indications of a central factor in uncontrolled and controlled shifts in cube perspective, this JOURNAL, 65, 1952, 89-90.
[2] Wolfgang Köhler, *Dynamics in Psychology,* 1940, 43-106.

by a 1-min. rest-period) were run under instructions to allow reversals to occur 'naturally.' S was told to say 'down' or 'up' (or 'neither' or 'both') to indicate changes in the cube's appearance. These reports were recorded on a kymograph by E. A second series, Series 2, with instructions to hold the 'up' phase, followed the first series, Series 1, by a rest-period of 1 min. and it consisted of two 2-min. trials separated by a rest-period.

Since the primary concern here was with the total *time* the 'up' phase was held in perspective, rather than the *rate* of reversal, the mathematical model involving square root transformations of reversal frequencies provided by Bruner, Postman, and Mosteller was not applicable.[3] In determining the relationship between S's intelligence and his ability to maintain the 'up' phase of a reversible cube under 'hold' instructions, it was desirable to take systematically into account the time it was in perspective under 'natural' instructions. One method of doing this, used by Mull, et al., was to express the time the 'up' phase was reported to be in perspective under 'hold' instructions as a percentage increase of the time the 'up' phase was in perspective under 'natural' instructions. Although this has the advantage of simplicity, it makes assumptions about the regression of time-scores in Series 2 on time-scores in Series 1 that cannot be substantiated in the calculations. By using this technique one assumes that a relationship exists between results from Series 1 and Series 2 without ever determining the nature of this relationship. For example, equal percentages based upon far different absolute scores would be treated equally, although the data might not justify this. To overcome this difficulty, analysis of covariance was used to control statistically the effects of the time-score under natural conditions in appraising the results under the control conditions. It revealed that there was no significant relationship between the two time-scores ($r = 0.06$).[4] A procedure for controlling time-scores in Series 1 while studying Series 2 was, therefore, neither necessary nor justified; all that was required was a consideration of the control time-scores as a simple variable.

When the total time the 'up' phase of the cube was held in perspective under 'hold' instructions was correlated with *ACE* total scores, the resulting coefficient was not significant ($r = 0.10$). In addition to this difference in the measurement of control scores, the discrepancy between these results

[3] J. G. Bruner, Leo Postman, and F. A. Mosteller, Note on the measurement of reversals of perspective, *Psychometrika*, 15, 1950, 63-72.
[4] The detailed summary of the analysis of covariance, together with descriptive statistics, etc., may be obtained by consulting the author's MS thesis which is on file in the Library of Purdue University, Lafayette, Indiana.

and those obtained by Mull et al. could also be the result of sampling differences. Mull et al., presumably used women as Ss while the Ss in the present study were men. The identical procedure was, therefore, repeated on a group of 31 women students. In addition to obtaining the time each S was able to hold the 'up' phase of a reversible cube, under instructions to do so for two 2-min. periods, similar data were obtained for the 'white' phase of a Köhler cross. Although the split-half reliability coefficients for the cube and the cross were fairly high (0.83 and 0.90 respectively), the correlations with ACE total scores were again not significantly different from zero (0.16 and $-$ 0.03).

These results suggest that caution be exercised and further evidence gathered before rejecting the null hypothesis that there is no relationship between intelligence and ability to control a reversible perspective.

The Menninger Foundation   DOUGLAS N. JACKSON
Topeka, Kansas

THIRTY-SIXTH ANNUAL MEETING OF THE WESTERN PSYCHOLOGICAL ASSOCIATION

The Western Psychological Association held its thirty-sixth annual meeting Thursday through Saturday, March 29-31, 1956, at the University of California in Berkeley. The total registration was 525. The program included 98 papers, 7 symposia, 2 films, 2 special luncheons, and the presidential address.

Allen Edwards gave the presidential address, entitled, "The social desirability variable in personality research," at the banquet Friday evening. Psi Chi and the Society for the Scientific Study of Religion arranged special luncheons. The California State Psychological Association held a business meeting Friday afternoon. The committee in charge of arrangements was under the chairmanship of Edwin E. Ghiselli of the University of California.

At the business meeting Friday afternoon the decision to meet at the University of Oregon in 1957 was reaffirmed. A location in or near San Francisco was favored for 1958.

Action was taken on the report of a special committee on the structure of the organization presented by Richard A. Littman, chairman. All recommendations for changes in the by-laws were adopted. The most important of these provided that the offices of secretary and treasurer be combined, but that a program chairman be appointed by the president to take over

the task of arranging the program and thus relieve the secretary of that responsibility. Other changes were designed to clarify the requirements for membership and make the system for election of officers work more smoothly.

In the mail ballot subsequent to the meeting, Leona E. Tyler was chosen President-Elect and David Cole Secretary-Treasurer. The President for the coming year is J. A. Gengerelli.

University of Oregon   LEONA E. TYLER

TWENTY-EIGHTH ANNUAL MEETING OF THE MIDWESTERN PSYCHOLOGICAL ASSOCIATION

The Midwestern Psychological Association met at the Hotel Jefferson, St. Louis, May 3, 4, and 5, 1956. Washington University was the host institution. Marion Bunch of Washington University served as Chairman of the Local Arrangements Committee. A total of 1064 persons registered during the meeting. The program consisted of 119 papers and 6 symposia. The symposia were concerned with the experimental analysis of verbal behavior, the first course in psychology, intuitive processes in clinical judgments, static and stationary models for human behavior, discrimination learning, and the prediction of adjustment. The presidential address, "Interference and forgetting," was delivered by Benton J. Underwood of Northwestern University. He showed that the forgetting of rote materials, which has been attributed to retroactive interference from learning taking place outside the laboratory, is probably due largely to proactive interference from previous laboratory experiences; when this proactive interference is removed, forgetting is 15% over 24 hours, compared to 75% shown in the classical studies.

W. K. Estes of Indiana University was elected president for 1956-57. New members of the Council are I. E. Farber, State University of Iowa, and C. H. Lawshe, Purdue University. With the election of 349 new members, the current membership is 1943. The 1957 meeting will be held May 2, 3, and 4 at the Hotel Sherman, Chicago.

University of Chicago   DONALD W. FISKE

FIFTEENTH ANNUAL MEETING OF THE CANADIAN PSYCHOLOGICAL ASSOCIATION

The fifteenth annual meeting of the Canadian Psychological Association was held in Ottawa at Chateau Laurier June 7 to 9, 1956. Total registra-

tion was close to 200. The program consisted of about 40 papers on a variety of subjects and 5 symposia. A session of films ran concurrently with the business meeting. Carroll L. Shartle of Ohio State University addressed an open meeting on "The functions and selection of administrators." The presidential address, "On transfer and human abilities," given by George A. Ferguson, was concerned with developing a generalized formulation of transfer, learning, and ability.

The new officers of CPA are: President, Julian M. Blackburn; Past President, George A. Ferguson; President-Elect, W. E. Blatz; and Secretary-Treasurer, Dalbir Bindra. The two new Directors are G. A. McMurray and F. R. Wake.

McGill University   DALBIR BINDRA

AN ACKNOWLEDGMENT

Robert Mearns Yerkes: 1876-1956

Few psychologists have contributed so much to psychology in so many different ways as Robert Mearns Yerkes, whose death from a weakened heart occurred on February 3, 1956, in his eightieth year. He was born on May 26, 1876.

First, of course, we think of Yerkes as the psychobiologist. It was his experimental contributions in that field while at Harvard that brought him the presidency of the American Psychological Association in 1917, an honor, as it happened, that coincided with the end of the first phase of his career. We can learn much about how he became a psychologist and about the early shaping of his character from the earnest and manly statement in his autobiography.[1]

Animals, domestic and wild, are the one recurrent theme in the three opening paragraphs of this autobiography. In particular, pet tortoises were to young Robert "treasures of entertainment," so it is not surprising that they later found a use among his first laboratory subjects. Life on a farm, however, can never be all fun for a small boy, not even on ancestral acres in the rich farming country north of Philadelphia. He tells us he became intractible, in particular did not get on with his father, almost ran away from home. Robert was shy and much alone, especially after the death of a younger sister to whom he was devoted, but he found an emotional anchor in his mother, and he delighted in companionship with seven of his grandparents and greatgrandparents. As a lad he learned to work, learned so well that during the rest of his life he never even thought of stopping. In a letter to the writer he confided late in his life, "During the greater part of my life I have habitually overtaxed my physical resources."

In youth, during a critical illness, he fell so readily under the influence of the family doctor that his interests turned toward medicine. A further nudge by a zoologist during his undergraduate years at Ursinus College, together with a generous loan from somebody else, put Robert where he was able to continue his education at Harvard (1897-1902). There he studied at first under a brilliant quartet of zoologists—Mark, Parker, Davenport and Castle—but later shifted to the group of philosophers and psychologists whose fame has become legendary—Royce, James, Palmer, Münsterberg and Santayana. "Great Teachers, Great Men, and Great Minds," he called them in gratitude.

---

[1] C. Murchison, ed., *History of Psychology in Autobiography*, 1932, II, 381-407.

While still a graduate student he began to publish a series of papers reporting experiments on the behavior of invertebrates. These, together with a considerable number of theoretical papers on comparative psychology, and his book *The Dancing Mouse* (1907), comprised the bulk of his output during his first professional decade. Through exceptional productivity, and with the particular support, we may infer, of Münsterberg and President Eliot, he early earned a foothold in the Harvard department.

When Yerkes first employed the term comparative psychology and sanctioned its use to designate the area of his appointment, it was in its then current sense of the psychology of all organisms excepting man. But at least as early as 1908 we find him preferring to divide the materials of psychology between animal and human psychology, thus reserving the term *comparative psychology* for "a method of investigation rather than a division of the field of psychology," a method that aims to trace developments or establish interrelationships.

Again and again Yerkes was to speak of his chief interest as the development of methods. In the preface to *The Dancing Mouse* he explained that he had used this animal as a means of exhibiting a variety of methods by which the behavior and intelligence of mammals may be studied.

During the next few years (1907-1917) he turned part of his attention to several methodological projects using human subjects, of which the best known was the development of the Point Scale (with Bridges, Hardwick, J. C. Foster and others), a method of measuring intelligence which he felt should supersede the Binet through the inherent superiority of a scaling (and therefore to him a comparative) method. Doggedly he promoted the Point Scale, which for various reasons, including insufficient norms and a lack of statistical sophistication, eventually dropped out of use. Yerkes also wrote a textbook, *Introduction to Psychology* (1911)—a "skeleton picture of the science of psychology," one that by his own prescription requires a manual of facts to supplement it.

In his autobiography he relates an incident that happened early in his youth to illustrate how, even then, he was aware of his preference for doing rather than knowing. An aunt had chided him for being unable to answer satisfactorily her impromptu quiz in geography and history. "When I admitted ignorance, she expressed surprise at the imperfection of my education. I well remember my mingled feelings of chagrin, rešentment, and disapproval, for her conception of education struck me as unsatisfactory. Even then my interest centered in constructive, creative effort toward the extension of knowledge, instead of in achievement of scholarship through mere accumulation of facts."

This distinction remained very real to him. His writings are never "accumulations of facts," and when, much later, he proposed, by way of exception, to make *The Great Apes* "an objective historical account of knowledge of anthropoid life," he confessed that "the manuscript tended to become increasingly critical and original as we proceeded with its preparation."

Throughout his Harvard years Yerkes was plainly fretting over the status of psychology as a science. "The current American psychology of today is a dismal mixture of physiology and psychology." It may be doubted whether any of the numerous papers he wrote on the relations between experience, physiology and behavior really satisfied him, for he came back to the subject again and again. But it seems that by 1910 he had come under Titchener's compelling influence and he wrote then in a scrupulous vein that is characteristic of him: "Although the greater part of my own contributions to science have been physiological, I feel that I am entitled by my experience in teaching psychology to include myself among the students of consciousness." In that same year he took the position in his *Introduction to Psychology* that the psychologist "should work, in the main, as though there were nothing in the world except psychic facts."

Obviously, then, he was unprepared to follow Watson, or later any of the other behaviorists. At that time the study of behavior was to him physiology. He wrote of Watson that "his is an 'illiberal attachment' to an assemblage of ideas which is in itself valuable, but which certainly does not monopolize the profitable possibilities of psychology or physiology." Looking ahead, we may note that it was not until the end of the twenties that he formally adopted the one distinctive designation for his field of research that he was completely happy about—"psychobiology." He made this term peculiarly his own. Psychobiologist he knew he was, and for the rest of his life basic terminology troubled him no more.

After 1910 Yerkes' future prospects at Harvard had become dimmed in consequence of a change of administrations. Eliot had retired and President Lowell "gently and tactfully" urged him to shift into educational psychology. Deeply identified with Harvard though he was, it would have been wholly out of character for Yerkes to yield to this suggestion, no matter how great the pressure. He was then eagerly at work, both in Cambridge and in his own field laboratory at his summer place in New Hampshire, on the development of his multiple-choice method for studying the intelligence of crows, rats and pigs. Most significant of all, a sabbatical half-year spent in Santa Barbara (1915) provided his first opportunity to experiment on the "ideational" behavior of monkeys and an orang-utan.

His writings from 1914 on make it clear beyond doubt that no other single life goal could be as important to him as the establishment of a laboratory in or near the tropics for psychobiological research on the anthropoids.

The year 1917 brought in quick succession two reorientations of Yerkes' life, one of them, however, turning out later to have been merely a threat to dislocate his life, but the other proving far-reaching in its consequences. After much deliberation he had just accepted an invitation to become head of the department of psychology at the University of Minnesota when America entered the World War. It was then April and the annual informal get-together of the group of experimental psychologists that pivoted around Titchener was in session at Harvard. This group named a committee with Yerkes, then president of the American Psychological Association, as chairman to express its conviction to the national association that in this emergency the services of psychology should be offered to the government. To that end the council of the American Psychological Association met almost at once, and, among many other appointments, made Yerkes chairman of a committee on the psychological examining of recruits. The fitness of this choice arose from Yerkes' development of the Point Scale, and his experience while at Harvard in clinical work in the Boston Psychopathic Hospital.

So it was that a new kind of opportunity to which he also proved to be by nature not unprepared to yield—though no one could have foreseen him in such a role—diverted him, at age forty, far from psychobiology into paths of highly responsible military planning and administration. During the war and afterwards for the remainder of a long life, his colleagues in psychology were to recognize in him a tireless champion of their interests. He undertook all sorts of tasks, some arising from formal responsibilities connected with the various capacities in which he served, and others, perhaps scarcely less numerous, self-imposed. Early and late, and with compulsive persistence, Yerkes fought to make psychology, as he would have said, "maximally useful."

His was the ardor and conviction of a missionary. If we should speculate on the origin of such a compelling drive, we might remember that his mother, whom he calls "the strongest influence in my early life," wanted him to be either a preacher or a missionary. After allowing for a change of direction consonant with his special interests and beliefs, can we not say that he did, in a sense, dedicate his life in that same spirit?

Happily Yerkes had the habit of writing full reports of his own activities and those of the men associated with him. Among these are the impressive and very full accounts of what he achieved and what he sought

to inspire others to undertake, in two simultaneous war roles—chief of the Division of Psychology in the Office of the Surgeon-General (at a rank no higher than Major!) and chairman of the Psychology Committee of the National Research Council, under whose auspices many psychologists applied their special proficiencies to military problems. In both these administrative capacities Yerkes was a pathbreaker without precedents to guide him. Working with utmost despatch, he and his committee assembled a stellar group of test psychologists to devise new instruments suitable for examining the intelligence of a very large number of recruits.

After this feat had been brilliantly accomplished, he turned to the mammoth task of convincing both military and civilian authorities that the new tests were not only practical but necessary, if army manpower was to be effectively distributed. What psychology owes to Yerkes becomes clear when we read in his autobiography that he struggled "to vanquish seemingly insuperable difficulties by overcoming the passive resistance of ignorance and the active opposition of jealousy, misinformation and honest disagreement." Yet in the end, 1,700,000 officers and men were tested. And then, humanitarian that he was, Yerkes could not help bemoaning the fact that, just as his psychological machine was perfected and authorization obtained to extend its coverage to the entire army, the war abruptly ended, thus sharply reducing the strictly scientific achievements that he had foreseen. So he felt—though he might have been reminded that for the moment he was forgetting his usual preference for method over facts, and that the new testing methods and their value had been triumphantly demonstrated.

When peace returned, Yerkes decided that he could not immediately abandon his role as a representative of psychology on the National Research Council in the city that he sensed was to become the center of organization for the nation's scientific life. The Minnesota position that had been held open for him he resigned, probably with few qualms, perhaps because the slate of appointments he recommended would create at Minnesota an almost wholly new department of which he could consider himself the godfather. At any rate, for the rest of his life he continued to take a keen and helpful interest in Minnesota's psychology.

In 1919, however, what Yerkes felt most strongly was the call to help maintain and increase psychology's momentum in Washington. It was well for his dream of establishing a station for primates that he did. His friendships with such influential scientists as Angell, Merriam, Hale, Dodge, McClung, Welch and Vaughan were probably in the end what turned the trick, and some of the immediate projects he set out to accomplish failed.

An anthropological and psychological study of the people of the United States foundered, probably because its grandiose sweep was far beyond anything warranted by the technical methods then available. Much of the ambitious and costly undertaking known, rather oddly, as a study of scientific problems of human migration, also petered out. Psychologists, nevertheless, should remember with gratitude that they owe promotion of their indispensable journal, *Psychological Abstracts,* to that period of the National Research Council's activity.

Most significant of all, however, was Dr. Yerkes' vigorous leadership as chairman of the N.R.C. Committee for Research in Problems of Sex, from 1922 to 1947. In their report, *Twenty-five Years of Sex Research,* Aberle and Corner describe the beginnings of this work: "A little group of earnest people were facing a vast realm of ignorance and half-knowledge, scarcely knowing even where or how to begin." Yet the same report informs us that both the National Research Council and the Rockefeller Foundation came to think of this committee's work as "an eminently successful example of research administration." The courage and social convictions of its chairman were largely responsible for this record. Through the years this small committee of wise men formulated their own policies and never lapsed into a mere doling-out agency for the funds they so successfully raised. In the very large amount of fundamental research they made possible, Yerkes justifiably took deep satisfaction, and perhaps he was most pleased that the committee gave Dr. Kinsey so much financial and moral support.

After five years of connection with the Research Council, Yerkes "escaped," as he wrote, "to more congenial activities." The psychologist Angell, then President of Yale, had set up in his university an Institute of Psychology in which Yerkes was appointed (1924) to a research professorship. Yerkes' first move was to visit a colony of captive primates in Cuba in order to size up the conditions needed for an anthropoid research station. The result was a temporary laboratory at Yale, where he demonstrated the feasibility of housing and conducting research on chimpanzees in New Haven.

This was all to be prologue. We have it on Yerkes' own word, however, that not even the old military challenges and discouragements equalled those he ran into in attempting to secure the permanent primate station. Undismayed and determined, during these first five years at Yale he turned out a book on chimpanzee behavior, three monographs on "the mind of a gorilla," reports of several research investigations, and, with his wife as co-author, the monumental treatise *The Great Apes, A Study of Anthro-*

poid Life. And we must not forget that then, as in every other period of his career, he was editing journals.

Finally in 1929, just in time to escape the blight of economic depression, full support for his plans was obtained. Yerkes became Professor of Psychobiology and first director of the Yale Laboratories of Primate Biology, which he described as "a primate Eden in Florida." Here at last chimpanzees could be bred and experimented with under nearly ideal conditions. It was the fulfillment of an idea going back to his days as a graduate student and one he had mentioned from time to time in his published writings. Yerkes' administrative headquarters were set up in the department of physiology at New Haven, where new laboratories of comparative psychobiology for research and instruction were also provided. Funds were available for field studies of primates in their natural habitats and several of these were of notable importance. Yerkes could now use his multiple choice method to study chimpanzee intelligence and insight. He carried on many other investigations and reported much naturalistic observation. Judged by any warranted standards, the productivity of the whole program was high both in quality and quantity, but it is doubtful that its indefatigable and modest director ever recognized how richly he had earned the many honors that came to him.

In 1942 Yerkes turned the directorship of the Orange Park laboratories over to Lashley, becoming director emeritus himself. Two years later he retired from Yale. The laboratories are now independently operated and bear his name: the Yerkes Laboratories of Primate Biology.

When the United States had become a participant in World War II Yerkes' attention focused on the consequences of the fact that for the first time a great part of the nation was to be harnessed for war. This expansion toward total war demanded all along the line stepped-up participation by psychologists. Planning for the future, he was sure, was most important of all. So presently by direct request he secured a mandate to form a committee on Survey and Planning for Psychology (actually a subcommittee of the N.R.C.'s Emergency Committee in Psychology). He was made its chairman and chose its members. Out of that committee's wide-ranging efforts and thinking came the idea, soon put into effect by an "intersociety constitutional convention" (1943), of restructuring the two large associations of American psychology into one "federated" society, thus ending the split between fundamental and applied psychology that had developed in the previous decade.

How Yerkes did love the committees he served on, and how hard he worked at them! For ten years more he kept busy hatching plans and urg-

ing that this and that be done. He drew up a "blueprint for American education"; he wrote on the place of psychology in the liberal arts curriculum; he served on the University Commission to Advise on the Future of Psychology at Harvard; he played the chief role in organizing a conference on military psychology; he was for years a member of the Scientific Advisory Committee of the Fels Research Institute in Yellow Springs, Ohio; and so on and on. In Henry Murray's phrase, he had a strong need for "ideo-dominance." He knew this fact about himself and admitted in his autobiography that he was "not a good follower."

The character of Robert Yerkes stands forth in all his life and accomplishments. Loyalty was the central fact about him: intense devotion to his wife, daughter and son, faithfulness to his many friends—mostly fellow scientists—and passion for service in the causes in which he believed.

Yet there was much more than this about him. Yerkes the scientist could become absorbed in laboratory research. Though a serious man, and one with little sense for either "fun" or banter, nobody could doubt his capacity for richly enjoying life. He speaks of his "joy in writing." His publications were voluminous, and his correspondence prodigious. As he put it, "I usually would rather write to a friend than eat my dinner." Those letters, besides talk of his work, his plans and his dreams, contained a deep tide of personal feeling that on most occasions ran far beneath the surface.

Late in life, Yerkes came to feel that his autobiography had been written too early. Because he could not do it over, he wrote a book-length manuscript, as yet unpublished, which he called informally his *Testament.* It ends with the statement of a *Personal Creed* that reads in part as follows,

I believe:
In knowledge of the natural order as basis of man's life.
In man's responsibility for his life, but not for eternity, destiny, immortality.
In the dignity and perfectibility of man as part of the natural order.
In usefulness through fellow service as incomparably worthy ambition.
In the priority of life over death, effort over prayer, knowledge over faith, and resolution over wishfulness.

RICHARD M. ELLIOTT

University of Minnesota

Warner Brown died February 6, 1956, after a very painful but gallantly borne illness.

He was born in Georgia on February 9, 1882. His grandfather Warner, a New Englander, "migrated west," first to New York and then to Georgia, where he was the owner of a large plantation. His own parents settled in Georgia and there began in Warner Brown that unique combination, which was his, of hard-headed Yankee honesty and skepticism, on the one hand, *and* of southern sensitivity of perception in the realms of the esthetic and the social, on the other. His entire education before college was through informal tutoring and his own prodigious and catholic reading. In his teens he read the Greek classics in Greek, the Latin classics in Latin, French literature in French, and had a wide acquaintance with English literature. He also acquired early an extensive knowledge of law and botany.

He moved with his parents to California in his late teens, entered the University of California in 1901, majored in philosophy under George H. Howison and Josiah Royce, and began his career in experimental psychology under George M. Stratton in the Psychology Laboratory, then the daring adjunct to the Philosophy Department. He received his A.B. in 1904, his M.A. in 1906, and went to Columbia where he received his Ph.D. in 1908 under Woodworth and where he was influenced by Woodbridge, Dewey and Cattell.

In 1908 he married Jessie Milliken, a botanist. They had two children —Frederick, whose death at seventeen was a difficult blow to absorb, and Ellen, now Associate Professor of Medicine at the University of California Medical School.

He became an instructor in 1908 at the University of California, Berkeley, where he served as teacher, guide, setter of standards, and relentless critic for forty-four years until his retirement in 1952. He was a wise and selfless department chairman for sixteen of these years.

In spite of his breadth of interests and knowledge, he was by temperament and conviction an experimental psychologist. He was a charter member of the Society for Experimental Psychologists. He thoroughly understood experimental design and was quick to see excellence or vital flaws in an experiment. He had an open-minded readiness to entertain new hypotheses but always sought to find "negative instances" in his own work and in that of others. He was much influenced in the early stages of his career by the rigorous thinking and method of Titchener, had a life-long

495

respect, admiration and affection for Woodworth, and in later years appreciated what the tight logic and the exciting consistency of Skinner's experimentation meant as a new approach. Few, if any, of his own publications deal with "theory" in any broad sense. He took "what appears before the eyes in a learning experiment as a basic datum," and showed an "unwillingness" to enthuse about learning theory, to paraphrase Skinner.

Several of Brown's early publications remain classics: his investigations of the probability of the perception of a difference in the region of the threshold; his careful experimentation showing that even very slight differences are discriminated with a certain frequency while large ones are sometimes not. The ogive curve of discrimination by small increments of difference is a quantitative, solid contribution of his. He later turned to investigation in the field of learning and memory. A study of reminiscence carried out in 1923 was characterized in 1953 by Osgood as one "which offered a very sober and reasonable" explanation of this effect. His work on whole-and-part methods in learning, on the retroactive inhibitory effects found in a card-sorting task, on the actual cues and orientations which humans build and depend upon in maze learning, form part of the facts in advanced texts on experimental psychology. He carried on and published studies in suggestibility.

His list of publications is relatively small, some thirty in number, aside from his text in *Introductory Psychology*. Why? He was such a perfectionist that many of his research endeavors, impressive to others, he would not publish, for he was his own most relentless critic and knew that many of his undertakings were non-definitive and not without flaw. He had a great respect for science and no tolerance for the many journal articles that he felt were carelessly done or prematurely published. Further, and perhaps equally important, he would get involved in new ideas that took all of his attention. He was more interested in investigation than communication; yet in a general introductory text, which he co-authored, he insisted on clarity and simplicity of presentation.

He spent a great deal of time duplicating and checking experiments which appeared in literature. In cases where he found "negative instances" from such work, it became known to students, who were then provided with unexpected insights into flaws in experimental design or unjustified conclusions drawn from experiments. In this way he sharpened their awareness of "good" and "bad" work.

Perhaps his most significant contribution to the department grew out of the fact that, whether one resisted it or not, he became the vigilant and incorruptible scientific conscience of its members. His approval was al-

ways sought and was never obtained for any scientific undertaking unless unequivocally merited by disciplined thinking and a neat performance. He was generous of his time and of his wide knowledge. He loved to teach and had a fine sense of pedagogical timing, encouraging the timid and deflating the pretentious. It can be truly said that he was more interested in making dedicated psychologists than in his own scientific achievements.

Let some of his students speak. Arthur I. Gates wrote in 1952 to Warner Brown—". . . for four years I enjoyed your ever-stimulating and incredibly generous guidance. . . . You needled me into tackling an investigation in my junior year, followed by others in each succeeding year. Most of these, including a sizable monograph, were published. . . . You were the most helpful, insightful and progressive teacher I ever had."

Calvin Hall: "Even now, some twenty years after leaving his direct surveillance, whenever I have an article ready for the press, I pretend I'm Warner Brown and read it through once more and always find myself blue-pencilling it drastically and muttering to myself, 'Tighten it up! Tighten it up!' "

John Gardner writes, "I believe that as a student, what impressed me most about Warner Brown was a quality of mind; and I think that the word for that quality is 'independent.' Some one said of Sainte-Beuve, 'he belonged to the great diocese; he was one of the independent seekers after truth.' If independent is the key word, the great diocese has never boasted more than a handful of parishioners. I think in a wholly unpretentious way Warner was one of that handful.

"What do I mean by independent? I mean that he wore no man's intellectual collar; he rode none of the band wagons of the day; he was free from stereotyped ideas and attitudes (almost as common among scientists as among others); he was unenthralled by the fashions and fads of the moment; he had an unhackneyed mind.

"He seemed to me to be committed to nothing beyond examining the evidence; and this he did cooly, incisively and astringently, with the eye of a man gifted to observe and appraise rather than to produce and communicate."

The stress placed on his critical professional temperament does not complete the picture of the man, for he was at the same time tolerant, kindly, and utterly democratic in his daily contact with people.

University of California

H. C. GILHOUSEN
JEAN W. MACFARLANE

BOOK REVIEWS

Edited by M. E. BITTERMAN, The Institute for Advanced Study

Theories of Perception and the Concept of Structure: A Review and a Critical Analysis With an Introduction to a Dynamic-Structural Theory of Behavior. By FLOYD H. ALLPORT. New York, John Wiley and Sons, 1955. Pp. xxii, 709.

Two purposes guided the author in this book: first, to provide a critical survey of theories about the nature of perception, and, second, to present his views about the problems which such theories should resolve.

In the first three chapters, Allport discusses the proper function of theories in psychology and in doing so covers familiar ground—how define "perception"? what status have sense-data? what are the criteria of objectivity? what is the logical nature of explanation? and so on. The number of philosophical questions raised is of course considerably greater than the number settled or even clarified. The chapter does enable Allport to give us an idea of his epistemological premises, and though these are matters which seem to have little bearing on the behavior of most psychologists, it seems a useful thing to know how a writer on perceptual theories stands with respect to them.

Next, Allport surveys a variety of theories about the nature of perception: the classical core-context theory, the gestalt theory, Lewin's topological theory, Hebb's cell-assembly theory, the Werner-Wapner sensory-tonic theory, Freeman's set theory, Helson's adaptation-level theory, Brunswik's functionalist theory, the Bruner-Postman hypothesis theory, and the cybernetic theories. These chapters constitute about 70% of the book and they will, I believe, be the most interesting to the general reader, since they provide a lucid and interesting survey of many ideas important to an understanding of perception.

In the concluding three chapters, Allport considers "the unsolved problem of meaning" and develops his concept of "event-structure." The various philosophical issues that enter this discussion are so numerous and so difficult, that a brief and accurate report of his opinions is out of the question. Instead I have chosen to say what I think he means by "the unsolved problem of meaning" and to explain how he thinks the concept of "event-structure" will help to solve it.

The proper analysis of the relation between non-verbal signs and what they signify is the fundamental problem which Allport regards as unsolved, but, before we can see this problem clearly, several semiotical distinctions are essential. We must understand first that the term "meaning" refers to a particular kind of relation among three terms, a *sign,* an *interpreter* of the sign, and a *significate* of the sign. The characteristic which distinguishes this triadic relation from all others is that one of the terms, the sign, has the property of being either true or false, and it is this characteristic of signs which provides the core of the psychological problem of meaning. The task of the psychologist is to provide an analysis of this complex relation in the language of psychology to the end that, given this analysis, we can discover empirical ways of investigating the sign-behavior of animals and men.

Now, in such an analysis, the sign will be identified as a special kind of stimulus, and the interpreter as an animal in a special condition (*e.g.* disposed to react in certain ways), but how can we identify the significate? When the sign is true, the significate could perhaps be identified by that thing or situation which makes the sign true. But what about when the sign is false? This case makes it clear that the significate cannot be regarded as a particular, but rather must be considered as a class of a certain kind. Consider the sentence "There is a sunflower in bloom at the North Pole." This sign signifies the class of blooming sunflowers in a certain geographic region. If the class has no members (*i.e.* is empty or *null*), the sign is false. If, however, there is at least one member of the class, the sign is true, and is said to *denote* that member. The same analysis can, of course, be made for any non-verbal sign, like a bell or light in a conditioning experiment, or a patterned card in a discriminative problem. This analysis enables us to make one additional important distinction. Classes can be identified and distinguished from each other in two ways. We may consider a class *in extension*, that is, in terms of its members, or *in intension*, in terms of the requirements that must be met if something is to be a member of the class. The class of American Nobel-Prize-winners-in-literature consists of Sinclair Lewis, Eugene O'Neil, Pearl Buck, William Faulkner, and Ernest Hemingway. When we have enumerated these individuals we have considered the class *in extension*. To be a member of this class, a person must (*1*) have been born in the U.S., (*2*) have met the requirements set by the Prize committee for work in literature, and (*3*) have been awarded the prize by a representative of that committee. Any person who meets all of these requirements is a member of the class, and when we consider the class in terms of these requirements, we are considering the class *in intension*. Considering then the significate of a sign as a class, we can clarify the difference between denoting and designating. A sign *denotes* the significate in extension, while it *designates* the significate in intension. The important point here is that two signs may have the same denotation and yet have different designations. The clearest instances are false signs which have the same denotation, namely, the null class, but which have different designations.

Now the problem for the psychologist is to provide an analysis of these complex relations which will permit one to determine the designations as well as the denotations of the signs which control behavior. Determining a sign's denotation is essentially like determining retention by a recognition test. The sign is presented with various potential members of the significate, and the interpreter of the sign then sorts out those that are members from those that are not.. But how can we determine the designation of a sign? Here is "the unsolved problem of meaning" and the point at which Allport introduces the concept of "event-structure."

Two alternative ways of conceiving of the designation of signs have been suggested in the past. First there is the attempt in the tradition of British empiricism to deal with such designation by means of images. The defining properties of the class (*i.e.* the significate) are given in the image aroused by the sign. Thus, the sign designates that class with members similar to the image. This treatment has the advantage of providing the kind of specificity of designation that the problem seems to require, but it has the disadvantage of introducing a mentalistic term ("image") as essential to the analysis. How can we determine the content of another's imagery, and what can we do with Kulpe's evidence of imageless thought? The failure to provide satis-

factory answers to these questions led to the decline of the imagery-doctrine. The other way of treating designation employs the CR instead of the image. Signs arouse fractional surrogates of the overt responses aroused by some US. The significate of a sign on this account consists in those stimuli which arouse URs similar to those aroused by the sign. The great advantage of this approach lies, of course, in its obvious objectivity. Its major disadvantage is its failure to provide the kind of designative specificity which the problem requires. How does the response of looking at a red wine differ from the response of looking at a white one? Or, for that matter, how does the drinking of red wine differ from the drinking of white? It seems quite unlikely that the specificity of designation characteristic of human behavior and that of the higher animals can reside completely in differences among responses. The afferent character of what is designated must somehow be incorporated in the analysis.

Before I turn to Allport's conception of "event-structure," I must make a confession to the reader. I proposed at the outset of this review to explain what Allport meant by the "unsolved problem of meaning," but what I have done instead is to say what I think he should have meant by it. My only excuse is that Allport's treatment of the problem is so thoroughly intertwined with his views on a host of philosophical issues that I found it impossible to disentangle the original question. When we come to his concluding chapter ("Outline of a General Theory of Event-Structure") the difficulty is even greater, for there the issues are represented on the canvas of the cosmos where the primeval forces of thermodynamics and gestalt ontology dominate the landscape. As a consequence, all that I can do to represent his views about event-structure is to express what I think might be the relation between what he seems to mean by "event-structure" and what I have described as the problem of meaning.

The concept of "event-structure" appears to be a way of regarding patterns of cortical events as mediators in the designation of signs. In other words, Allport appears to suggest that, through a knowledge of the qualitative structural properties of cortical processes, we may be able to find a basis for class-designation which has the specificity that images are reputed to have without their concomitant subjectivity. In fact, he seems to say that, when we have begun to specify the structural properties of such processes, we shall be able to derive implications both about physiology and about behavior which will be directly testable. In this sense, perhaps, the concept of "event-structure" solves "the unsolved problem of meaning."

Does all this mean that Allport has coined a new term for whatever was named by the phrase "cognitive structure?" I do not think so. What Allport appears to have done is to specify some of the spatial and temporal characteristics which cortical events must have if they are to play their proper role in sign-behavior. Hebb was the first to return our attention to this problem, while Allport is the first to press the analysis to the biological level, on the one hand, and to the cultural level, on the other.

What will interest the reader, however, is not the key to the problem which Allport offers, but the many locks which he attempts to open with his key. His survey of the theoretical problems in perception is clear and exciting, and I suspect that his book will be valued primarily as an excellent description of the present status of perceptual theory. At the end of Allport's book the reader is left with

the feeling that he has been on a long trip, in which he has seen so many exciting things that he can only remember one thing clearly—how much more pleasant it would have been with less baggage.

University of California

BENBOW F. RITCHIE

Schools of Psychoanalytic Thought: An Exposition, Critique, and Attempt at Integration. By RUTH L. MUNROE. New York, Dryden Press, 1955. Pp. xvi, 670.

Ruth Munroe defined a comprehensive set of objectives for herself in writing this book. The first was to present, for a wide reading audience, a sympathetic, semi-technical exposition of the major schools of thought in psychoanalysis. The second was to compare the various theories and to evaluate each critically. The third was to offer a possible approach, framework, or structure within which the different positions could be organized and integrated. The various schools of thought are considered in three separate sections of the book. One is devoted to Freud and the "Freudians." It includes not only a description of the basic position of Freud, but also the elaborations of Hartmann, Melanie Klein, Eriksen, Kardiner, Anna Freud, and others. A second section presents the ideas of Adler, Horney, Fromm, and Sullivan, and a third is devoted to the work of Jung and Rank.

These expository sections are for the most part very well done. They are clearly written, in a style that is oftentimes conversational. One receives the impression that the author is truly trying to 'feel' her way into the material and, with a positive attitude, is trying to present it without the interference of a strong bias. Some measure of preference is reflected, however, in the amount of elaboration afforded each of the views. Despite the general sympathetic style, there is a progressive change from the rich, meaty treatment of the "Freudians" to the rather meager, skeletal presentation of the views of Jung and Rank. A more serious criticism can be made of the plan of organization used by the author for the section on the "ego" or "self" theorists (Alder, Horney, Fromm, and Sullivan) in relation to that for the libido-theorists (Freud and the others). The utilization of five homologous chapter headings in each section is particularly suited to a smooth-flowing offering of the libido-position, but hardly seems like the best medium for the non-libido group. In general, it forces constant cross-comparisons among the views of the ego-theorists which interrupt the continuity of the presentation and make for excessive repetition. In the interest of clarity, this reviewer would have preferred the ego-theories to be treated separately, like those of Jung and Rank.

Comparison and criticism are handled in a novel way, being confined to the close of each chapter, clearly labeled, in double columns, and in smaller type. Two general categories are stressed in critical evaluation. One deals with the important aspects of personality neglected by each school, and it is here that the author shines. She is able to call not only upon psychoanalytic literature, but also upon a wide knowledge of her own specialty (psychology), to fill in the gaps. Especially cogent for the points of view she considers is her insistence on the importance of the "nonsexual drive systems" (hunger, motility, and so forth) in the development of the organism and its sense of "selfhood." The other major area of criticism is concerned with what the author calls "reductionism," by which she means overextension of a theory to account for phenomena that are much better explained by recourse to other principles. In the critical portions of the book, numerous instances are cited in which either theories

have been over-generalized to account for a wide variety of facts, or rich clinical observation has lost some of its luster by being reduced to single, universal, causal categories. While all psychoanalytic schools fall prey to this disease, Munroe finds Freud and Sullivan least susceptible. Their positions are described as "multi-dimensional," a positive value-term which indicates that many variables in these systems have been delineated and their dynamic relationship to one another considered. Jung also is included in the above category insofar as he has described the various aspects of the self, but his concepts are criticized as referring to fixed universals rather than to dynamic, changing systems.

The criticisms are brought together in the book's final section. Here a cogent critique of the concept of aggression is offered in which the "single instinctive drive" of the Freudians is found to be inadequate and various aspects of this complex phenomenon are discussed. It is in this section also that the author collects a good many of her scattered comments on the integrating framework for psychoanalytic schools. This "view of systems," as it is called, is in itself contentless; it is similar in nature to the structural views proposed by Lewin, Murphy, and Angyal. Munroe conceives of the organism as a series of operating systems which can be classified in a variety of ways. Some of these systems are rooted almost solely in the structure of the human body, e.g. the vegetative systems. Some, like the sensory systems, are much more dependent on factors outside the organism for their initiation and operation, while others, like the drive-systems, have an intermediate character. Some systems by their relative permanence may take on the appearance of universals, while others are more fleeting in the life-history. To study the intrinsic operation of these systems and their interactions is, for the author, the most fruitful way for a psychological science to proceed.

The reviewer had some small hope that the last section of the book would go beyond a reiteration of the preference for such an approach and deal with its intricacies instead. This hope was not realized fully, but the bait cast out was sufficiently promising to cause him to look forward to a future effort by the author.

There have been several attempts in the last twenty years to present the kind of material that Munroe has given us here. Judged in the light of these previous efforts, as well as for its intrinsic value, this is a very fine book. It is clear, lucid, and informative. Its biases are openly expressed and for the most part held in check, particularly in the expository sections. Freudian content is used by the author as a sort of 'home base' and occasionally as a yardstick against which to measure the positive contributions of the other schools, but by no means is this the polemic of a blind disciple. The perspective of the author is oriented toward a marriage of the findings of psychoanalysis within the framework of modern psychology.

Princeton University

IRVING E. ALEXANDER

Small Groups: Studies in Social Interaction. Edited by A. PAUL HARE, EDGAR F. BORGATTA, and ROBERT F. BALES. New York, Alfred A. Knopf, 1955. Pp. xv, 666.

This book contains 55 theoretical and empirical selections from the literature on small groups and a valuable annotated bibliography of 584 items. The readings are organized in to three parts: (1) historical and theoretical background, (2) the individual in social situations, and (3) the group as a system of social interaction.

The historical and theoretical papers are assembled into three chapters: Early Theory (Durkheim, Simmel, Cooley, and Mead); Early Research (Terman, F. H. Allport, Riddle, Thrasher, and Turner); and Current Theory (eleven papers on a miscellaneous set of conceptual and empirical problems). The early theory is sociological, the early research mainly psychological, and the current theory half and half (or perhaps neither according to the traditional designation of these fields). Included under current theory are several treatments of such limited aspects of groups as leadership or communication, some of which take the form of miniature conceptual systems. Another interrelated cluster of papers deals with "systems of interaction," a point of view which has been advanced by Parsons and Bales, and which is further documented by a rather large number of selections in the last part of the book. Finally, there are a few extracts from the writings of "significant theorists" which outline theoretical orientations rather than theories and which do not do justice to the theorists.

The second part, dealing with the individual in social situations, contains three loosely related chapters. The first deals with the classical psychological topic of "together and apart" but, unfortunately, contains nothing that summarizes the now extensive body of accumulated findings. The second, on social perception, includes four papers which lay out in as comprehensive a way as could be expected in such small space the currently more interesting problems and issues. The third chapter is on "consistency of the individual." The eight studies reported here include some most imaginative work, but this reviewer is unable to determine in what sense the chapter-title ties them together. Considering this part as a whole, one's reactions are mixed: most of the individual papers report significant and interesting research, but all together they illuminate only a small segment of the area designated by the heading. Perhaps a topic more specifically related to small groups would have served better to organize these selections.

The third part, on the group as a system of social interaction, is the heart of the book. It contains four chapters which consider communication-networks, interaction and equilibrium, role-differentiation, and leadership. The research which is used to elucidate these topics was conducted almost entirely on groups of people engaged in discussion. While this restriction limits rather seriously the range of safe generalization, it does result in an excellent coverage of what is known about conferences and problem-solving groups. Even within this narrower range, however, the reader (or teacher) will himself have to organize the findings into a comprehensive set of propositions, but the richness of the material will make such work rewarding. The view of the field presented by these selections is eclectic, though not complete. Eight of the studies were conducted directly under the influnce of Bales and constitute an excellent source of information about this development in the work on small groups.

Assuming that this volume represents a reasonably accurate sampling of the research on small groups, what may be said to characterize this field? Three features seem most pronounced.

First, the field is relatively new and rapidly expanding. The oldest items included in the book are a theoretical piece by Durkheim (1902) and an empirical study by Terman (1904). The earliest empirical study reported in the bibliography is one by Triplett which appeared in this JOURNAL in 1898. Aside from the 'historical'

papers, the earliest study reprinted appeared first in 1938, and well over half appeared after 1951. The production of research since World War II has been impressive indeed.

Second, the field of small groups is an offspring of mixed parentage. While several disciplines may claim relationship, the parents are undoubtedly psychology and sociology. The infant itself is not sure of its personal identity. Those doing research in the field may call themselves "psychologist," "sociologist," or "social psychologist," but it is hard to tell one from another without a professional directory. In the interest of neatness it would be satisfying to limit psychology to the study of the individual and sociology to the study of the group, and something of the sort was attempted in the distinction between the second and third parts of this book, but the compartmentalization simply will not work. Individual behavior and development are inextricably interwoven with the properties of groups, and these very properties are themselves made up of behavior and the interrelations among individuals. Apparently this state of affairs produces tension within all those working in the field. Perhaps a creative reorganization of the problem may soon be expected.

Finally, from a theoretical standpoint, the field is poorly organized. In view of the sudden and recent upsurge of empirical work and the mixed theoretical heritage of the research, it is not surprising that integrative conceptions should be scarce. Most of the studies included in this volume make use of the experimental method and are designed to test (or at least to demonstrate) hypotheses, but the fact remains that few of the hypotheses stem from comprehensive theory. The most impressive theoretical work achieves its rigor by restricting its attention to a very few variables and, up to the present time, different theorists have considered different sets of variables. It is to be hoped that before long these may be brought together into more comprehensive systems without sacrificing the rigor of the component parts.

The book serves the field well by bringing together in a convenient place many publications which otherwise would be hard to consider side by side. Its bibliography is a definite aid. Its index is virtually useless because the topical headings are so broad that there are literally dozens of items under a single, undifferentiated heading. The value of the collection as a textbook will depend upon the preferences of the individual teacher.

University of Michigan   DORWIN CARTWRIGHT

Without the Chrysanthemum and the Sword: A Study of the Attitudes of Youth in Post-War Japan. By JEAN STOETZEL. New York, Columbia University Press, 1955. Pp. 334.

"Given the right approach, no culture is incomprehensible to a mind formed in any other culture" is the assumption of Jean Stoetzel, a French social psychologist, who, with a Dutch linguistic scholar, made up the Unesco research team commissioned in 1952 to discover the post-war attitudes of Japanese youth. His "right approach" involved the use of such familiar psychological tools as the representative sampling poll, open interviews, and projectives, a modified *TAT* and the Allport-Gillespie Autobiography to 2000 A.D. That so much could be accomplished by this team in a few months, *e.g.* the polling of 2,671 people, was due to the coöperation of Japanese professors and the experienced Japanese National Public Opinion Research Institute.

As the title suggests, the book affords a marked contrast with Benedict's

Chrysanthemum and the Sword. Through library sources, interviews with Japanese in this country during the war, and comparative anthropology, Benedict built up a simple monolithic model of Japanese cultural behavior governed by the hierarchical obligations of *on* and *giri*. Stoetzel, on the contrary, contends that "In complex societies with a history and a past" there "must be many-faceted models," and that "the analysis of form must be abandoned in favor of quantitative evaluations." This greater attention to individual differences is especially important when old and new patterns of behavior are in competition. Japanese critics and foreign visitors had found it difficult to locate anywhere a reality corresponding to the rigid model of Benedict; at best they had thought it the ideal of a former subgroup, the *samaurai* or warrior caste. Stoetzel's results show variability of response on all questions.

The Unesco inquiry in Japan (along with similar studies made at the same time in India and Germany) was based on the assumption that if the youth were found self-reliant, realistic, and secure, then democracy would be well guarded in the future. The 23 questions of the poll were in these areas: attitude toward social change, attitude toward authority, readiness to take part in political life, interest in and attitude toward foreign countries, and personal values. From the poll and the projective tests, Stoetzel found Japanese youth ambivalent and with rather definite indication of passivity, insecurity unless supported by others' advice and even presence, and such escapism as ignoring of politics and excessive addiction to movies. He sums it up: "What the young Japanese of today will be tomorrow probably depends less on their present attitudes and personality traits than on the path to which the Japanese society commits itself; and they themselves are still too immature and show too much dependence for us to be able to think that they will have much say in the fresh impulse . . . that will be given to that society with the end of the Occupation."

This conclusion obviously cannot be supported without comparative data from other countries and long-term studies. For example, would a similar survey of American youth show them politically active and strongly determining the direction of our society? Stoetzel recognizes that interviewing methods may tap only superficial aspects of personality. The large-scale Rorschach studies still in progress by George DeVos in Nagoya and surrounding villages, with Japanese clinicians trained by him in American technique and interpretation, will provide a crucial test of the concepts of both Benedict and Stoetzel. Other difficulties which hindered Stoetzel came from the well-known ambiguity of the Japanese language; several of the Allport-Gillespie questions were misunderstood. Moreover, the law forbidding the national polling organization from undertaking any police or partisan activity prevented a frontal attack on ideological questions.

Both for its theory and data this book has undoubted value for the social psychologist and the student of personality. The appendix of nearly 100 pages gives the statistical and sociological particulars of the sample, a complete tally of even the open-ended questions, two protocols of the ten plates of the modified *TAT*, eight autobiographies of the future, and an annotated bibliography of Japanese studies of youth. This completeness should make comparisons with other cultures both easy and rewarding. Most of all this book ought to be on the required reading list of the State Department.

University of New Hampshire   GEORGE M. HASLERUD

Studies in Communication. By VARIOUS AUTHORS. London, Martin Secker & Warburg, 1955. Pp. vii, 182.

If there were a journal devoted to communication, this book could be one issue, and, if all the issues were as good, it would be a very fine journal indeed. Since there is no single journal devoted to the subject, the only media that seem to be open for communication about communication are symposia and their frequent off-spring, published symposia. The present volume is constructed out of a series of lectures sponsored by the Communications Research Centre of University College, London, and it illustrates nicely the dilemma of this young field of inquiry. Its strength is its weakness; the topic of communication is an excellent rallying point in the struggle against academic compartmentalism, it provides issues that challenge the Renaissance Man in all of us, but, just to the extent that it does flow over the entire landscape, there is no single vessel that can contain it all. Where else can philosophers, biologists, physicians, engineers, art critics, philologists, phoneticians, and teachers of composition collaborate except in such a volume as this?

For example, T. B. L. Webster has contributed a very interesting discussion of the problems that the ancient Greeks faced in communicating with one another, how their language changed to support their philosophy and poetry, and the criteria they used to discipline their reasoning. Psychologists would be especially interested in the detailed history of the term *psyche* from Homer to Demokritos. This scholarly analysis obviously has some connection with J. B. S. Haldane's brief survey of communication within (neurophysiology and endocrinology) and between (ethology and ecology) organisms, but what is it? Are these together in the same book because of an overly free association of terms, or is there really some basic unity behind the *-sis* suffix in Greek and the dance of the honeybee? We need an ideational centrifuge to separate out what does from what does not belong here; perhaps that is part of the purpose of books such as this.

From a purely personal viewpoint, I was sorry that no psychologists were asked to participate. There are continual references to psychological problems: A. J. Ayer does not want to deny to psychologists the right to investigate imageless thoughts; C. Cherry feels that because psychologists are 'tightly coupled' with their observed system they may speak of probabilities as degrees of belief rather than as relative frequencies; G. Vickers observes that the 'kick and carrot' psychology of the nineteenth century is now discredited in economic analyses; J. Z. Young regrets that the most important part of human nervous activity is handed over to the various schools of psychology; R. Wittkower says that Gestalt psychologists and psycho-analysts cannot explain why similar symptoms may be found in art of the highest order as well as in abominable vulgarizations. These and many similar remarks might have come out somewhat differently if a psychologist had been asked to comment on them; or perhaps the psychologist would have come out differently. At any rate, the contributions seem to suffer at several points from misconceptions of modern psychology.

In his introduction, B. Ifor Evans promises us that this is but the first volume in a series; it will be followed by other volumes on the results of investigations and discussions which are at present being pursued. We can hopefully anticipate that the new Communications Research Centre will so crystallize present enthusiasm into

future accomplishments that we may all see more clearly what integration of these diverse departments is possible.

Harvard University   GEORGE A. MILLER

Suicide in London: An Ecological Study. By PETER SAINSBURY. London, Chapman and Hall, 1955. Pp. 116.

According to Sainsbury, the problem of suicide is that of assessing the relative contributions of the factors causing conflict and tension within the personality. Among the environmental determinants of suicide, a critical situation, or series of situations, may so disrupt a person's customary means of satisfying his needs that suicide may be attempted. In light of this formulation, two investigations were undertaken: "I. A statistical correlation of suicide rates in the twenty-eight Metropolitan Boroughs and the City [of London] with selected indices of their social characteristics. II. An analysis of social and other information pertaining to 409 suicides reported to the coroner for North London during a three-year period."

The "London boroughs showed: (a) Significant differences between their suicide rates, and a remarkable consistency in rank order of rate over three decades in spite of considerable changes in the composition of their populations. (b) A significant correlation of suicide rates with rates for the following characteristics: social isolation (e.g., persons living alone, and in boarding-houses); social mobility (e.g., daily turnover of population, and number of immigrants); and two of the indices of social organization (divorce and illegitimacy). As regards socio-economic status, suicide tended to increase in the middle class and decrease with poverty. . . . Unemployment and overcrowding rates showed no correlation with suicides" (p. 90).

"The composition of the population of the boroughs by age, sex and marital status did not explain the distribution of suicides: on the contrary, the factors which . . . account for the ecology of suicide in London may also afford an explanation of the suicide rates peculiar to these groups. Twenty-five percent of suicides in London had some abnormality of personality. . . . The external stresses that were thought to have precipitated suicide were socially derived in a third of the cases. Physical illness was a factor in a quarter. Both tuberculosis and cancer were more common among the suicides than among the population at large. . . . Mental illness was the principal cause of suicide in a third of the cases. The incidence of senile and arteriosclerotic psychoses (post-mortem findings supporting the latter diagnosis) was remarkably high (10%)" (p. 92).

Suicide is an unequivocal and reportable human action, but it remains a psychological puzzle. Kallmann has shown that it is not genetically determined. In two of every three suicides there is no evidence of physical or mental pathology. The accounts given by persons who have made real, but unsuccessful, suicidal attempts add little to our understanding of the motivation which culminated in attempt. If there exists little or no difference between individuals with respect to biological potentiality to suicide, if age, sex, and economic status cannot be shown to be true determinants of suicide, then why do more suicides occur in one urban section than in another? Can one argue that architecture, number of persons housed per residential room, similarity or dissimilarity of interest, aptitude, or education, really act differentially to produce significant differences in the incidence of suicide? Somehow

these "statistically significant" relationships remain possibilities rather than causal associations. Until a satisfactory explanation can be evolved which will account for the demonstrated fact that occurrence of suicide is approximately twice as frequent in May than it is in July or August the ecological differences which Sainsbury has demonstrated seem comparatively small.

Psychiatric Institute, Columbia University   CARNEY LANDIS

The Psychology of Thought and Judgment. By DONALD M. JOHNSON. New York, Harper and Brothers, 1955. Pp. x, 515.

This volume is really two books, one of which deals with *thought* and the other with *judgment.* Both may properly be described as learned and exhaustive. The treatment of judgment summarizes the extensive literature concisely and critically; that of thought rather brashly reflects the current experimental and theoretical uncertainty.

The activity of problem-solving is analyzed into three main components—preparation (after Graham Wallas), production, and judgment. "A person may be said to have a problem if he is motivated toward a goal and his first goal-directed response is unrewarding." A *problem* is indeed a difficult thing to specify. The reviewer has made an attempt, not very successfully, and knows the extreme caution necessary. It does not seem, however, that Johnson's definition will do; the *detour* of Köhler would, for example, be cut out, and the word "unrewarding" is unsatisfactory. Stated in general terms, the mechanism of thinking, according to Johnson, seems to be that the "motivation for solving a problem" will "capture" the set of the solver, directing attention to the problem-situation. This does not seem any great advance over Watt's formulation of exactly 50 years ago, especially if combined with that of Ach at the same time. A kind of précis of their statement would be that a task is accepted, from which spring *determining tendencies* uniting the task with the problem-data. Little seems to be added by the use of the word *set,* the vagueness of which has been proclaimed vigorously by Gibson. This similarity to an almost forgotten formulation is not necessarily a fault; the facts may indeed demand such an analysis. It is interesting to notice that, just as the concept of *determining tendencies* was overworked by Ach, the intensely modernistic Johnson has somewhat overworked the concept of *set.* There is, of course much more experimental evidence on the latter than Ach was able to bring forward. The whole of this part of the book is a valiant and somewhat prolix attempt which does not seem quite to come off. It ought to be read by psychologists, however, if only to throw into relief the difficulties of dealing with this critical problem. No better treatment exists, and the faults are those of present day formulations. What is said about the relation of learning to thinking is valuable.

Chapters 9-11 deal primarily with *judgment.* Here the writer seems to come suddenly to his own stamping-ground and gives a valuable survey. Judgment has for many years been in itself a recalcitrant problem. We do not understand it yet, but the process is used many times a day in every laboratory, assisted though it may be by modern statistical and other techniques. The scale-maker and the psychophysicist will find the treatment in these chapters worth reading.

The book should be available, for refeernce at least, to every psychologist.

University of Oxford   GEORGE HUMPHREY

Psychological Statistics. By QUINN MCNEMAR. Second Edition. New York, John Wiley & Sons, 1955. Pp. vii, 408.

The format, level of difficulty, and general approach of the second edition are essentially the same as the first. The revisions consist chiefly of additions of new material and extended discussions of topics previously covered; the only deletions are of an occasional nonessential phrase or sentence. The author and publishers call attention to these changes and additions: an expansion of the treatment of statistical inference and the theory of hypothesis-testing (including some discussion of the Neyman-Pearson principles), a progression from large-sample to small-sample methodology for continuous variables, a binomial approach to chi-square, the Cochran method for handling several correlated proportions, the exact-probability method for fourfold tables, additional methods for comparing variabilities, a detailed presentation of the underlying models for the analysis of variance and their implication for the proper error-term, a discussion of the limitations of the latin square (à la McNemar's article in the *Psychological Bulletin* of 1951), and a brief chapter on distribution-free methods—only the sign, median, and Mann-Whitney *U* tests are included.

Considered with respect to additions and changes *per se,* the revised edition is almost without exception superior to the first, both in clarity of exposition and in the inclusion of valuable additional material. New developments in statistical methodology and the growing literature on the application of existing techniques to psychological problems confront the author of a statistical textbook for psychologists with difficult decisions regarding choice of subject-matter. What one views as sins of commission or omission in this revised edition will depend on the purpose for which the book is being considered.

McNemar writes in his preface, "The widespread adoption of the first edition of this textbook suggests that it has been found useful in introductory courses although it was not written primarily for that level." The second edition makes more concessions to the beginning student, but the discussion of some elementary topics, *e.g.* random sampling, remains inadequate. I believe that the second edition, like the first, will be most useful as an intermediate textbook. An equally explicit exposition of the models for analysis of variance, particularly in relation to typical experimental designs in psychology, is not readily available elsewhere. The value of the book as an intermediate text would have been increased by material on transformations for the analysis of variance, an extension of the treatment of the relative efficiency of statistics and designs, illustrations of the nonparametric tests cited, and more discussion of their advantages and disadvantages.

University of California   DOROTHY H. EICHORN

Einführung in die Medizinische Psychologie. By GEORG DESTUNIS. Berlin, Walter de Gruyter, 1955. Pp. ii, 218.

This *Introduction to Medical Psychology* consists of two main parts, entitled respectively "Personality Psychology in Medicine" (structure of personality, fourteen drives, motivation, noetic functions, psychodiagnostics, typology) and "Theory of Neurosis" (including a section each on psychosomatic illness and organic syndromes). A third part, ten pages in length and entitled "Personality and Illness," is appended. All conceivable names of syndromes are mentioned; the symptoms are

occasionally described in brief, but never systematically discussed. Most psychopathology results from overstrictness or overindulgence of parents, though the basic causes are necessarily constitutional and even phylogenetic (*e.g.* sexual sadism and masochism in the spider). The reader is informed that psychoanalysis has contributed something, provided it is stripped of its exaggerations and distortions—but no attempt is made to analyze such contributions and to define specific areas of misplaced emphasis. The treatment is rounded out by nineteen photographic portraits of patients, *e.g.* "an aggressive-depressive personality," "an epileptic psychopath," "an enuretic child," "a sticky epileptic," "a stutterer," "a homosexual transvestist magnificently adorned."

According to the author's foreword, his book was written as a text for practicing physicians and students of medicine and psychology. Unfortunately this reviewer is unfamiliar with the more recent German-language publications in the field. It may well be that one or more good clinical texts exist which can be read with profit by those for whom the book under discussion was intended. In any case, Gordon Allport's *Personality* is now available in translation. Meili's *Psychologische Diagnostik* gives an excellent introduction to the field of psychological testing and diagnosis, and there are a number of widely known classical discussions of neurosis which, with their varying systematic points of view, make excellent reading and stimulate critical discussion. While perhaps not all reach Allport's level of literacy, none of them (like the work at hand) refers to Dougall, Mc.

College of Medicine, University of Illinois   MARIANNE L. SIMMEL

Becoming. By GORDON W. ALLPORT. New Haven, Yale University Press, 1955. Pp. viii, 106.

Allport has long been a master of stating an important issue with economy and clarity, taking up a firm position with respect to it, and then defending that position with force and wit. In this book he has done it again.

Based on the Terry Lectures at Yale, this small volume is a summary of Allport's philosophical outlook as well as of his thought on various persistent issues in psychological theory. He has given no ground since 1937. The arguments for uniqueness, wholeness, active intention, contemporaneous and complex motivation, high-level integration, cognitive dynamisms—things which are now grouped under the "Leibnitzian tradition"—are as forceful as ever, and the old enemies, positivism, operationism, elementarism, geneticism, equivalance of species, environmentalism— the "Lockean tradition"—are punished with the same old weapons.

From the point of view of the psychologist, the book does not represent Allport at his best. The trouble springs, it would appear, from the nature of his assignment: "that of assimilating and interpreting his discipline as it relates to human welfare and to religion broadly conceived." The assignment called for condensation, of heroic proportions, before an audience that must have been made up largely of people other than psychologists. In consequence, as it seems, there are too many places in which the stand on an issue seems to proceed more from the over-all philosophical position than from a hard-headed confrontation of the facts; there are too many straw-men around, too many stereotypes, and too many over-simplications for the sake of a neat point. The sallies against the positivists, behaviorists, Freudians, and other

"Lockeans" are not as good-natured, as delightfully meaningful to the psychologist, as one would have hoped.

What makes this reviewer particularly nervous is his impression that when the chips were down Allport's pragmatic humanism might prove stronger than his value for science. The Lockean position is to be rejected, it seems, because in trimming down the image of man (has it done so, really?) it has bad consequences, such as, perhaps, favoring totalitarianism (which, of course it does not). What if the Lockean position should turn out to be the truest one? What makes this reviewer a little ill-humored is the feeling of having been preached to, of having been told what to do about his sins, without his having been convinced that he has sinned in the first place. Yet insofar as Allport set himself the task of stating and defending a philosophical position, one that embraces an ideology of psychology, he has succeeded admirably. His statement is bold, lucid, scholarly, internally consistent. Most psychologists probably will find here much to disagree with, much to be aroused by. It would be a good thing if some of them were provoked to prepare comparable statements of their own.

University of California   NEVITT SANFORD

The Subnormal Mind. By CYRIL BURT. Third edition. London, Oxford University Press, 1955. Pp. xix, 391.

This new edition of Burt's well-known book leaves us still admiring its author but disappointed that so little new has been incorporated. The field of subnormality has been undergoing rapid recent evolution in standpoints, terminology, research advances, and new outlooks—so much so that any new book runs the risk of being outmoded by the delays of publication unless the most current trends have been incorporated.

Perhaps one is misled by the title. The author is at some pains in the preface to defend his originally "forward look." He rightly deprecates the divorce of mind from body, and of different facets of "mind" from each other; but this does not lead to entitling the book "The Subnormal *Child*," which would more effectively represent the integration principle, and one is left confused by later categorization which seems to disintegrate mind from behavior, and behavior patterns from individuality. (The separate chapters deal with the mentally deficient, the dull or backward, the delinquent, the neurotic, and with the asthenic and the sthenic neuroses.)

The author is undoubtedly familiar with the children who compose these groups, but dynamic process gets lost in descriptive traits, and the overlaps which confuse clinicians are little emphasized. If one might say that the clinical psychological management of children with deviations of learning and adjustment has been moving from differential diagnosis to therapeutic (dynamically oriented) synthesis, the issue becomes more apparent.

If one views the book's title literally, one expects major emphasis on the second and third chapters which deal with "the mentally deficient" and "the dull or backward." One also expects some contradistinction by the very separation as well as by current (in the U.S.A.) trend of psycho-socio-educational versus essentially medical management. That we are confused in this country regarding these distinctions is painfully evident in our terminology, child-evaluation procedures, and management-

programs. For example, 'mental retardation' includes sometimes (*1*) below the mental test-norm used, (*2*) a deviation multiple of such subnormality (which varies according to test used, age of child, and other variables), sometimes a general area, other times a restricted area, (*3*) clinical mental deficiency, (*4*) specific defects or disabilities, and so on. Sometimes mental retardation is an assumed equivalent for scholastic retardation, general or specific, and, when 'mental deficiency' is contrasted with 'dull' or 'backward' normality, the discrimination is (here) made usually on the clinical distinctions of symptom-complex, battery psychometry, history, etiology, prognosis. These issues are reflected in the relatively newer terms of 'educable' versus 'trainable' (in California 'point 1' and 'point 2' respectively).

Burt urges that "mental deficiency is a legal rather than a psychological term," but it is also an extra-legal clinical term (*cf.* Tredgold). He notes that the criteria in the Mental Deficiency Act turn on social adaptability and in the Education Act on educational capacity (p. 63) and he concludes (p. 65) that the term "must not be taken to describe any well-defined clinical entity," but this appears to be belied by subsequent discussion of clinical issues, *e.g.* causes, ascertainment, incidence.

The concept of "morally defective" is hardly more convincing than of "intellectually defective" and both seem inconsistent with the basic standpoint of holistic psycho-biology advocated elsewhere in the book. Actually these confusions pervade the field of research as well as practice and create a semantic barrier to sound communication. The above observations thus reflect this reviewer's disappointment that the ideologies are not resolved in this edition. Certainly, few men are better qualified than Burt to do so.

Bellingham, Washington　　　　　　　　　　　　　　　　　EDGAR A. DOLL

Expression of the Emotions in Man and Animals. By CHARLES DARWIN. New York, Philosophical Library, 1955. Pp. xi, 372.

A new edition marred by a puerile preface (in which Margaret Mead enthuses over the notion that the word *communication* be substituted for Darwin's *expression*) and 10 pages of added illustrations (drawings of angry and fearful dogs by Lorenz, snapshots of Balinese doings by Bateson, and photographic studies, made at the University of Louisville's 1954 Conference on Culture and Communication, of unquestionably communicative, if not cultured, participants busily building the "science of kinesics"—"there is a team situation in which the sophisticated photographer is part of the research team and the scientists who discuss the problem themselves become subjects who illustrate the behavior which they are studying").

Dublin　　　　　　　　　　　　　　　　　　　　　　　T. S. KILLARNEY

THE ORIENTATION OF FORMS ON THE RETINA
AND IN THE ENVIRONMENT

By IRVIN ROCK, New School for Social Research

The appearance of a form is, in part, dependent upon its orientation, despite the fact that the proximal stimulus, the image of the form, remains unchanged insofar as the relationships of its parts to one another is concerned. Everyday life observation tells us that recognition will become more difficult or fail completely when the orientation of certain forms deviates to any considerable extent from the normal one. This, of course, is only true for forms which have one particular normal orientation as, for example, text and script or human faces. It is not true for forms which are seen in all orientations as, for example, keys or boxes. There is some experimental evidence dealing with this problem. Dearborn, and Gibson and Robinson have shown that familiar forms will be correctly identified more frequently when they are shown in their original (or usual) orientations than when they are shown in orientations displaced by 90° and 180°.[1]

The term 'normal orientation' is, however, ambiguous because it can be defined in two different ways. A form may be regarded as normally oriented if it maintains its customary orientation in relation to the upright direction of the environment or it may be regarded as normally oriented if it maintains its customary orientation in relation to the perceiver. The first case implies that the form, even with the head tilted, would remain normal or upright; the second, that a form whose image con-

tinues to fall in the same retinal orientation remains normal or 'upright' even if it ceases to be upright in the environment.[2] This can happen when one turns the form so as to preserve its customary relation to the head as, for example, when one turns a book to read while lying on one's side in bed. When the head is tilted the two types of uprightness or normalcy no longer coincide.

Common experience in situations like the one of reading in bed suggests that only when orientation in the second or retinal sense is changed do we have difficulty in recognition. We have trouble reading if we hold the book evironmentally upright when our head is tilted 90° to one side; conversely we do not seem to have any trouble when we turn the book parallel to the tilted head although now the book is no longer in its normal environmental orientation.[3] Köhler, in distinguishing between these two possible meanings of orientation, comes to a similar conclusion on the basis of two simple observations.[4] In one, the observer bends his head over and looks between his own legs at a photograph of a person held in an upright position by an assistant. The picture then looks quite strange and unrecognizable although it continues to appear environmentally upright. In the second, the observer continues to look from his inverted position, but now the picture is turned objectively upside down. The picture then looks normal in all respects though inverted in the environment.

There are, however, two studies which contain some evidence that the orientation of a form relative to the environmental upright is not without effect on recognition although on the whole they both support the above conclusions. (1) In a recent study, Thouless experimented with a reversible puzzle picture which could be seen as a bearded person in one orientation or as a sailor in an orientation reversed by 180°.[5] The observer had to invert the position of his head, i. e. look at the ambiguous figure from between his legs, and say which of the two faces he saw. Thouless found that 12 of 14 adults and 11 of 13 children saw the objectively inverted (retinally upright) face. The remaining subjects (2 adults and 2 children) saw the objectively upright (retinally inverted) face. Results of a control experiment indicate that normal orientation of these forms with respect to the environment was exerting an influence in the observations of the latter subjects. (2) In an earlier study Oetjen tested S's ability to recognize nonsense syllables and forms shown under different orientations of the syllables and forms, or of S's body, or of both, i. e. either S or the exposure-materials remained upright, or one or the other or both were tilted 90°.[6] The results indicated that recognition was very good only

---

[2] For a given orientation of an image on the retina we assume that there is a corresponding orientation of the pattern of excitation in the visual cortex. Henceforth 'retinal' will be used without further reference to presumed conditions within the nervous system.

[3] Of course one might think that the crucial factor in the case of text or script is eye-movement since we are accustomed to moving the eyes from the left to the right side of the head in reading and, therefore, unless the book is tilted with the head, the customary eye-movements would be inappropriate. This does not explain, however, why fixation of even a single word of script will produce the same results.

[4] Wolfgang Köhler, *Dynamics in Psychology,* 1940, 15-30.

[5] Robert Thouless, The experience of "upright" and "upside-down" in looking at pictures, *Miscell. Psychol. Albert Michotte,* 1947, 130.

[6] Friedrich Oetjen, Die Bedeutung der Orientierung des Lesestoffs für das Lesen und der Orientierung von sinnlosen Formen für das Wiedererkennen derselben, *Z. Psychol.,* 71, 1915, 321-355.

when forms preserved their customary orientation with respect to the observer. With retinal orientation of forms held constant, there was, however, a very slight but consistent superiority when the forms were upright in the environment.

The weight of evidence seems to support the view that of the two logically separate ways of defining the term "orientation," retinal orientation is of primary importance. This evidence, based as it is on the use of familiar and very complex materials, may, however, represent special cases. Further evidence is needed.

The experiments reported below were undertaken in an effort to explore this possibility. It may be well to point out that no attempt will be made to study in detail the question of exactly how any particular figure is changed in appearance as a result of disorientation. Rather it will be assumed for the present that forms do, in general, look different in different orientations and the question of which meaning of orientation is the more important will be investigated.

EXPERIMENTS I AND II: AMBIGUOUS FIGURES

The technique employed in the first series of experiments, Experiments I and II, is similar to the one used by Thouless in that figures are used which are ambiguous and present different axes for the alternative modes of organization. The main difference in the figures used here is that the axes of the two possible ways of organizing the figure are only 90° apart, not 180° as in the Thouless experiment. Two figures were created with this in mind. They are shown, Fig. 1, in a neutral position where the two possibilities can best be seen by turning the page 45° to the left and to the right. Fig. 1A will then be seen as an outline map of the U.S. when turned to the left and as a profile of the face of a bearded man when turned to the right. Fig. 1B (modified from Gibson and Robinson) will be seen as a Thurber-like outline of a dog when turned to the left and as a profile of a chef when turned to the right.[7]

Experiment I: Procedure. O is seated in a chair in front of a table at the far end of a corridor. He tilts his head 90° by resting it on its side on a foam rubber mat on a table. He is then instructed to bring his head as far forward as possible until his eye is in front of an aperture in a board through which he looks toward the far end of the corridor.[8]

O is told that he will be shown pictures on a small cardboard screen within

---

[7] Gibson and Robinson, *op. cit.,* 39-47.
[8] The purpose of the long corridor and the small aperture through which O views the scene relates to an experiment to be described in which O remains upright and the scene is viewed through a prism which causes it to appear tilted. The long corridor provided an excellent visual scene. To keep conditions more or less equivalent in all other respects, O here views the same scene through an aperture.
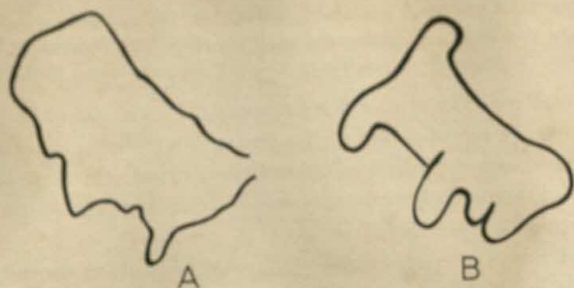
FIG. 1. AMBIGUOUS FIGURES USED IN EXPERIMENTS I AND II

the scene and that he is to identify what he sees as accurately as possible. A drawing of one of the ambiguous figures 7 in. in diameter is placed on a stand 12 ft. from O. It is shown briefly by lifting a cardboard shield. It is shown in such a way that one of its axes of organization is also tilted 90° to the same side as the head. It, therefore, remains retinally normal. The axis of the other mode of organization remains environmentally upright. This is schematically illustrated in Fig. 2 for head left (Condition I) and head right (Condition II). The question the experiment seeks to answer is this: Which organization will be achieved more frequently, the one which remains normal in orientation on the retina or the one which remains upright in the evironment?

A ready signal is given just before each presentation. Occasionally, a figure had to be shown more than once to an O before one or the other organization was perceived. This, however, in no way interferes with the validity of the results since,
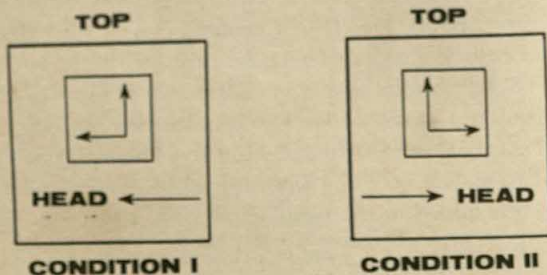
FIG. 2. SCHEMATIC ILLUSTRATION OF THE POSITION OF O's HEAD
DURING EXPERIMENTAL CONDITIONS I AND II

so long as the O remains naïve as to the ambiguity of the figures, the crucial question is which way is the figure seen under the particular experimental conditions.

Two conditions are used. In Condition I, O's head is tilted 90° to the left. In Condition II, O's head is tilted 90° to the right. The introduction of the two conditions serves the purpose of balancing the aspect of each figure which is upright in the environment. The order of presentation of the two figures is varied for different Os. Approximately 20 Os participated in each condition.

Results. The results are shown in Table I. The environmentally upright organization is indicated by the number in italics. The cases where no response was made as well as cases where it was questionable from the written response exactly what was seen are eliminated from the table since these are of no importance for our purposes. This explains why the amounts in the table do not add up to the number of *O*s used, and why in this and the subsequent experiments the number of *O*s used is given in approximate terms. Percentages are calculated from the totals of those who did report seeing the figure one way or the other. That organization of the figure whose axis remained environmentally upright was seen in the great ma-

TABLE I

FREQUENCY OF RECOGNITIONS OF TWO MODES OF ORGANIZATION
IN EXPERIMENT I: HEAD TILTED 90°

(Italics indicate environmentally upright figure.)

Condition	Fig. 1 A		Fig. 1 B	
	map	profile	dog	chef
I. Head left	*15*	1	*13*	5
II. Head right	5	*12*	1	*16*

jority of cases. Totaling both conditions for both figures, there were 56 selections of an environmentally upright organization (or 82%) and only 12 of a retinally normal organization (or 18%).

After combining responses to the two figures and for the two conditions, any difference between environmentally and retinally upright aspects cannot be due to inherent preferences for one or the other organization in each figure. Chi-square for this comparison is, of course, highly significant. Actually, in designing the figures in the first place, it was thought desirable if the two alternatives were about equally good. To achieve this, the figures were modified a number of times and pretested. It was possible to obtain figures that were fairly evenly balanced in ambiguity.

The results thus show a strong preference for the environmentally upright organization. It should be pointed out that if the results had indicated an equal choice for the two organizations of a figure it would have meant that both determinants were exerting an influence and of about equal strength. This is so because neither figure was shown in an orientation that was neutral with respect to either factor. In the figures as shown, one alternative is always retinally normal and the other is 90° displaced. But with the presence of the environmentally upright alternative in each figure, the retinally normal alternative is not able to dominate.

Experiment II: Procedure. A test of the influence of the retinally normal factor

alone is given by the results of Experiment II, in which the figures are presented in the neutral position shown in Fig. 1 with *O*'s head tilted 45°. A headrest is used which keeps *O*'s head at the desired angle. Under these circumstances, one axis of organization is retinally normal, one is retinally tilted by 90°, but neither is favored as far as orientation in the environment is concerned. Otherwise the procedure is identical with that of Experiment I.

Results. The results for 20 *O*s are quite clear. The retinally normal organization is seen in the majority of cases (29 of 38 responses, or 71%) and the preference is significantly different from chance at the 1% level. Hence it may be concluded that this would also be the case in Experiment I were it not for the overriding effect of the environmental orientation of each figure. The role of the latter is surprisingly great in Experiment I. The different retinal orientations at best play only a minor role.

Thus the results show not only that environmental uprightness plays a role in recognition but that it may predominate when the two factors are placed in opposition to each other. This contradicts certain everyday life experiences and certainly contradicts previous experimental findings referred to above. The question therefore arises whether these results can be expected only with a procedure such as the one employed. Is the selection expressed by the *O*s genuinely based upon the fact that normal orientation in the environment is more important for *recognition* than normal orientation on the retina? Or is it perhaps spuriously based upon the fact that *O* has an *expectation* that he will be shown something environmentally upright (or conversely the absence of an expectation that he will be shown anything environmentally tilted). If the latter were true this experiment would not be very conclusive. A technique was sought which would not suffer from this possible limitation. The experiments described below were undertaken with this in mind.

EXPERIMENTS III-VIII: SIMILAR FIGURES

In the present series of experiments (Experiments III-VIII), *O* is first shown a nonsense figure in a particular orientation while he is in an upright posture. Following this, a separation between the retinal vertical and the environmental upright is brought about either by tilting the scene or by tilting *O*'s head. *O* is then shown a card containing six figures, two of which are identical with the original. One is now made retinally normal in the same orientation to *O*'s head as during the initial exposure. The other is made environmentally upright as during the initial exposure. The question is, which of these will *O* recognize as the one shown previously. Four other figures on the test card serve to prevent *O* from immediately seeing that there are two figures both of which are alike except for differ-

ence in orientation. For the same reason all *Os* are naïve concerning the presence of the two possibilities on the test card. Otherwise a choice between them would become something of an intellectual task.

O is told to make his choice rather quickly and spontaneously upon surveying the different figures. It is assumed he sees all the figures and selects the one which is phenomenally most similar to the figure he remembers. This experiment meets the objection, concerning expectation, that might be raised to the ambiguous figures referred to above. Now it is not a question of which of two *alternative* organizations will emerge but which of two forms, both of which are perceived, looks most like the one seen previously. This technique is, therefore, more like the situations in daily life in which we actually encounter forms in novel orientations.

Experiment III: Procedure. The scene is tilted 45° by means of a Dove prism. *O*'s head remains upright. One of the critical test-figures is retinally normal and the other appears to be upright within the now tilted scene.[9] Prior to placing his head inside a box containing a prism at the front, *O* is shown the figure he is later to identify from among the six choices in the test card. Three such training figures are used and they are shown in their 'normal' orientation together with their corresponding test cards in Fig. 3. *O* is asked to inspect each training figure for 15 sec. with the instruction to familiarize himself with it so that he will be able to recognize it later on. Following this they are shown it again for 5 sec. in reversed order.

Fig. 3A is somewhat like a Chinese character. Fig. 3B is a closed nonsense figure. Fig. 3C is structurally similar to a letter of the alphabet. These figures were chosen to be more or less representative of both open and closed figures. Each figure is drawn in india ink on a white card $4\frac{3}{4} \times 7$ in.

O then places his head inside the box so that he looks out at the scene of the corridor described earlier. The test cards containing the six figures are shown after he views the tilted scene for a minute or two. The test card for each figure consists of drawings of the six figures arranged in a circular pattern. Each drawing is in india ink on a circular white piece of cardboard about $2\frac{3}{4}$ in. in diameter mounted on a white cardboard 13 in. sq. In this way the orientation of either of the critical figures could easily be changed for subsequent experiments. Adjacent to each of the six drawings is a colored spot to provide a means for *O* to identify his choice to *E*. Numbers or letters could not be used since they would have to be oriented one way or another and as a result might influence *O*'s choice. Before the test begins *O* is told that the next card to be shown contains one of the three figures shown previously. He is instructed to pick out the one which looks most like it. He is then shown the three test cards, the order of presentation being systematically varied.

[9] This experiment is predicated on the assumption that the vertical axis of the perceived scene will, to a very great extent, serve to define the environmental upright direction; cf. S. E. Asch and H. A. Witkin, Studies in space orientation: II. *J. exp. Psychol.,* 38, 1948, 455-477. An equivelent procedure is one in which *O* tilts his head 45° and one critical figure remains upright and one is tilted 45° so as to become retinally normal. It has been found that both procedures give about the same results.

The cards are shown at a distance of 12 ft. Occasionally O noted the two possibilities in a given test card. Sometimes when this happened they had a clear preference. When they did not prefer one or the other they were scored as 'both' and their response for this card does not enter into the totals.

There are two variations of this experiment. In Condition I, the scene is tilted 45° to the left. The test cards in Fig. 3 show how the cards look for this condition
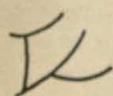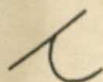
A B C
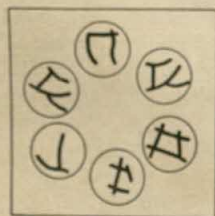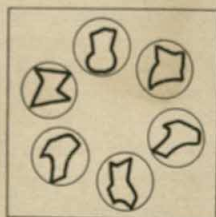
TEST CARD A   TEST CARD B   TEST CARD C
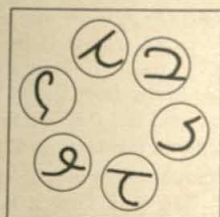
FIG. 3. NONSENSE FIGURES AND TEST CARDS USED FOR COMPARISON

in that one of the two critical figures in each card is tilted 45° to the right. When the scene is tilted 45° to the left by the prism, this figure becomes retinally upright. The other critical figure in each card is upright with respect to the card and the scene. It maintains its relation to the scene when the entire view is tilted by the prism. In Condition II, the scene is tilted 45° to the right. For this, the one critical figure in each card is simply tilted 45° to the left instead of the right. The other critical figure remains as is. The position of the two critical figures in the circular pattern on the test-card is shown in Fig. 3. Although the position of these figures on any one card is not changed, when the scene is tilted in the opposite direction in Condition II, their position with respect to O automatically changes. Also, the position of the critical figures was not the same for the three test-cards nor were the colors used to designate these figures the same. About 20 Os were used under each Condition.

Results. The results given in Table II show the number of times the Ss selected the critical figure upright in the scene (U) or retinally normal (RN) in the scene. As in Experiment I, responses scored as both or neither are eliminated.

It can be seen that the figure which appears to be upright (U) is selected somewhat more frequently than the one which remains retinally normal (RN). In the case of Fig. 3C, the dominance is marked. Altogether, for the three figures combined, there were 63 (62%) selections of the environmentally normal (U) and 38 (38%) selections of the retinally normal (RN). This difference yields a chi-square which is significant at the 2% level and means that the preference for the environmentally upright (U) over the retinally normal (RN) cannot be explained as a chance occurrence.

TABLE II

FREQUENCY OF CHOICES OF TWO CRITICAL FIGURES IN EXPERIMENT III:
SCENE TILTED 45°

Condition	3 A		3 B		3 C		All	
	U	RN	U	RN	U	RN	U	RN
I. Scene left	11	7	12	6	14	7	37	20
II. Scene right	7	8	7	8	12	2	26	18
Total	18	15	19	14	26	9	63	38
Percentage	55	45	58	42	74	26	62	38

These same Os, while viewing the tilted scene, were also shown the ambiguous figures used in Experiment I. One alternative of the figure is upright with respect to the room as now seen and hence is tilted 45° from its normal retinal orientation. The other is also tilted 45° from its normal retinally upright position but to the opposite side; hence it is tilted 90° from the upright axis of the perceived scene. In other words the figures are presented exactly as shown in Fig. 1 except that the tilted scene is aligned with one or the other alternative. Almost all recognitions were of the organization upright with respect to the scene. This result is cited as further evidence of the tendency to favor figures which appear upright within the tilted scene although in this case there is no direct conflict with the retinal factor. Neither organization is favored as far as retinal orientation is concerned.

Experiment IV. Since a tilted visual scene alone has an effect on recognition, the question arises whether phenomenal uprightness as such is the crucial factor favoring the figure upright in the scene. Perhaps what is important is not that a figure appear environmentally upright but only that it preserve its orientation in relation to the immediately surrounding visual framework.

Kopfermann has shown an effect of this latter kind where a square appears as a diamond or conversely a diamond appears as a square when surrounded by a 45°

tilted strongly." We might explain this effect in a manner parallel with the present experiment if we assume that when seen normally the sides of a square are parallel to the vertical and horizontal lines of the scene. Hence a tilted square will look like a square again if a frame of reference is tilted so as to be parallel to the sides of the figure. In other words, since environmental orientation is a factor in maintaining phenomenal similarity of form, it may only require the relationship of the lines of the figure to a few external lines such as the vertical and horizontal lines of the scene. If this is true, then it ought to be possible to obtain results similar to those we have obtained by using not the entire visual scene but only a small surrounding frame of reference as in the Kopfermann demonstration. The present experiment was designed to test this possibility.

Procedure. Instead of tilting the entire scene, only the test cards containing the six figures were tilted. They were attached to the end of a small stand at an angle of 45°. The scene, viewed from an upright position through an aperture, was equivalent in all respects but tilt to the one provided by the prism in Experiment III. Tilting the card has the result that one of the critical figures, normal in its relation to the sides of the card, retains the same relation to the sides of the card as when the card is not tilted. It is, however, retinally tilted by 45°. The other critical figure becomes retinally normal when the card is tilted, just as it does as a consequence of the prism in Experiment III. In this experiment, however, this figure is also upright with respect to the remainder of the scene which surrounds the tilted card. In other respects the procedure was like that used in Experiment III. There were two conditions (Conditions I and II) in which the cards were tilted to the left and right respectively. Approximately 20 Os were used in Condition I and about 15 in Condition II.

Results. The results of Os selections are given in Table III. It will be seen that the retinally normal figure (RN) is chosen more frequently in every

TABLE III

FREQUENCY OF CHOICES OF TWO CRITICAL FIGURES IN EXPERIMENT IV;
TEST CARD TILTED 45°

Condition	3 A		3 B		3 C		All	
	U	RN	U	RN	U	RN	U	RN
I. Card right	6	30	5	30	3	13	14	33
II. Card left	5	18	10	11	6	17	21	44
Total	11	18	15	11	9	18	35	77
Percentage	38	72	45	68	34	76	31	69

case. Altogether there were 77 responses to this figure and only 35 to the one which remains upright with respect to the card. This difference is significant at the 1% level. Hence there is a reversal of the results obtained in Experiment III where the figure which appeared upright in the scene was heavily favored. Since the only difference between these two experiments is the

*** Hertha Kopfermann, Psychologische Untersuchungen über die Wirkung zweidimensionaler Darstellungen körperlicher Gebilde, *Psychol. Forsch.*, 13, 1930, 293-364.

tilt of the immediately surrounding field instead of the entire scene, it is quite clear that the relationship of the figures to the visual lines of a frame of reference is not a sufficient factor. It is true that in this experiment there is a larger frame of reference provided by the lines of the scene of the corridor, and this larger frame is in agreement with the retinally normal figure. If the crucial factor were the relationship of the figure to a frame of reference one would, however, expect the immediately surrounding one to exert the stronger influence.[14] Therefore we must conclude that the important factor in Experiment III is the appearance of the figures as phenomenally upright or phenomenally tilted in the environment.

Since the results of Experiment III show a role of environmental orientation where the latter is achieved by a tilted scene alone, they may be interpreted as demonstrating in a somewhat novel way the role of a visual frame of reference for perception of the upright. After all, the critical figure which remains retinally normal is, in this case, also upright with respect to gravity. Hence were it not for the fact that the tilted scene exerted a very strong influence in determining the upright direction of objects within it there would be no reason at all to select the critical figure which is upright within the scene but tilted 45° from the gravitational upright. That this is an effect of the tilted scene on the phenomenal upright and not just an effect of surrounding visual lines of reference is clear from the negative results of Experiment IV. Aside from this fact, however it should be made clear that this study concerns itself with the problem of orientation in form perception and not with the problem of perception of the upright. Since environmental uprightness is one of the two factors under investigation, however, factors which influence perception of the upright (frame of reference, gravity) also enter into our considerations.

Experiment V. In Experiment III, the critical figures were only 45° apart; in the present experiment a 90° difference is used.

Procedure. In this experiment O bends his head sideways 90° resting it on a table as in Experiment I. He looks through an aperture in a board at the scene of the corridor. Except for this difference the procedure is like that used in Experiment III. One of the critical figures remains upright in the scene. The other is parallel to the head and therefore is tilted by the same amount as the head. In Condition I, O's head was tilted 90° to the right; in Condition II, 90° to the left. Approximately 15 Os were used in each condition, most of whom also participated in Experiment I.

Results. The results are given in Table IV. In every case the environ-

[14] Karl Duncker, *Ueber induzierte Bewegung, Psychol. Forsch.,* 12, 1929, 180-259; Asch and Witkin, *op. cit.,* 455-477.

mentally normal figure (*U*) was selected more frequently. Altogether there were 66 selections of the environmentally normal figure and 27 selections of the retinally normal figure. This distribution differs significantly beyond the 1% level from chance expectation that the two factors, retinal and environmental, are equally effective. It is desirable, however, to compare these results with those of a control experiment (Experiment VI) before their meaning can be fully understood.

Experiment VI. Experiment VI was performed to determine to what extent the retinal normalcy of one figure accounts for the number of in-

TABLE IV

FREQUENCY OF CHOICES OF TWO CRITICAL FIGURES IN EXPERIMENT V:
HEAD TILTED 90°

Condition	3 A		3 B		3 C		All	
	U	RN	U	RN	U	RN	U	RN
I. Head right	12	4	12	4	12	5	36	13
II. Head left	12	4	9	4	9	6	30	14
Total	24	8	21	8	21	11	66	27
Percentage	75	25	72	28	66	34	71	29

stances in which it is selected in Experiment V. Would it be selected less frequently if *O*'s head were not parallel to its axis, thereby no longer affording it retinal normalcy?

Procedure. O's head is now tilted 45°, midway between the two critical figures, under conditions otherwise identical with those of Experiment V. Under these circumstances, neither figure is favored retinally (unlike the situation in Experiment V but one is favored because it is environmentally upright whereas the other is not because it is tilted 90°. The 45° head-tilt is accomplished by clamping the box containing the board with the aperture at a 45° angle to the table. Thus when O's head is aligned with the box it is of necessity also tilted by 45°. Approximately 30 Os served in this experiment.

Results. There were 77 (83%) selections of the environmentally upright (*U*) figure and 16 (17%) selections of the environmentally tilted figure. As was to be expected, the upright figures are selected most of the time. Comparing these results with those of Experiment V, we see that the environmentally upright (*U*) figures are selected somewhat more frequently (83% instead of 71%). The chi-square comparing the two distributions is 3.65 which just misses being significant at the 5% level.[12] We may conclude, therefore, that the factor of normal retinal orientation accounts to some extent for the somewhat greater number of selections of the

---

[12] Here the direction of differences is predicted, hence we may be warranted in concluding that the two distributions differ significantly.

90° tilted figure in Experiment V than in the present experiment where neither figure is favored from the standpoint of retinal orientation.

The two experiments (V and VI) taken together show that both factors contribute to the results of Experiment V. Since the environmentally upright figure is chosen more often in both experiments, the role of this factor is clear. With the environmental factor the same and with neither favored retinally, the effect is even larger. The results of Experiment V are surprising because they not only show that environmental uprightness is effective but this factor is more powerful here than in Experiment III with the 45° tilted scene. From experience in daily life, it seems very clear that a 90° retinal disorientation produces greater difficulty for recognition than does a 45° retinal disorientation. Therefore one might have expected that in Experiment V the 90° retinal disorientation of the environmentally upright critical figure would make for greater difficulty. If it did, it was outweighed by the even greater difficulty in recognizing the other figure because of its 90° tilt in the environment.

Experiment VII. An experiment will now be described in which the same conditions as those of Experiment V were employed in order to investigate the following problem: The results of Experiment III and IV taken together demonstrated (1) that an important consideration when the scene is tilted is the change in the direction of the upright and (2) that one determinant of the upright, the visual frame of reference, in isolation from the other, gravity, suffices to influence the appearance of forms. Would the reverse be true? If phenomenal uprightness is indeed the important factor here it is plausible to assume that gravity alone would play an important role in influencing the appearance of forms; but if gravity alone can have this effect it means that the visual lines of the scene need not be present at all.

The following observations suggest that this may be the case. If one forms an after-image of a square and tilts one's head 45°, then with the *eyes closed,* the after-image can be seen in a field without visual directions. The square now appears as a diamond. Conversely, the after-image of a diamond (a 45°-tilted square) appears as a square with the head tilted 45°. The orientation of the figure in relation to the environmental upright as determined only by gravity suffices to produce the effect. Experiment VII was undertaken to confirm this reasoning by eliminating a visual frame of reference altogether.

Procedure. In Experiment VII the scene which O perceives through an aperature is one in which all lines of direction have been eliminated. The test cards are presented behind a circular opening cut in a large white cardboard which extends in all directions beyond the edges of the field of view allowed by the aperture. The edges of the test card are not visible, but only the six figures arranged in the circular pattern as shown in Fig. 3. Hence, although the field is illuminated, a visual

frame of reference is eliminated as effectively as in a completely darkened room such as is generally used to achieve this end.[18] By lifting briefly a circular cardboard shield, the appropriate test card behind it is exposed just as in all the previously described experiments. Care is taken so to arrange the lighting that no cues are provided as to the direction of the upright. In other respects the procedure follows that used in Experiment V. O's head is tilted 90°.

In a preliminary experiment, one critical figure was tilted 90° so as to be parallel to O's head and the other remained upright, just as in Experiment V. The results did not reveal a very strong superiority of the upright version. We had realized, however, that the procedure might not be correct in one respect. With a field without clear visual directions, such as we have here, we should expect, when the head is tilted, to find a constant error in the direction perceived as upright. For a 90° head-tilt the Aubert phenomenon is known to occur. This means that the upright may be displaced by as much as 20° or more in the direction opposite to that in which the head is tilted.[19] As a matter of fact the Aubert phenomenon did occur. Each O first had to judge when a small metal rod seen in the center of the circular cardboard appeared to be upright. All Os showed the expected error. (This fact may be offered as proof that the procedure did successfully eliminate all visual lines of reference.) This meant, however, that the one critical figure, which was objectively upright, did not *appear* to be upright. Consequently, the procedure was revised by first determining each O's subjective upright with the rod, and then pre-setting the upright versions on each test card accordingly. The retinally normal versions remained tilted 90°. In Condition I, O's head is tilted to the right; in Condition II, to the left. Ten Os were employed in each condition.

Results. The results are given in Table V in terms of whether O selects the critical figure phenomenally upright (U) or the one which remains retinally normal (RN). Altogether, there were 39 (75%) selections of the figure upright in the environment and 13 (25%) selections of the figure retinally normal. This difference is significant beyond the 1% level. The results answer the question posed in the affirmative. As a matter of fact the results are quite similar to those of Experiment V where the visual scene was present. It may be concluded, therefore, that a form's appearance is influenced by its apparent orientation in the environment and that it does not matter which cues serve to establish its orientation for this effect to occur.

Experiment VIII. The last experiment in this series is one involving a 180° difference between the two critical figures. Köhler's little experiment makes it clear, at least in the case of a form like the human face, that it is retinal and not environmental reversal that makes for the greatest change.

---

[18] See Witkin and Asch, Studies in space orientation: III. Perception of the upright in the absence of a visual field, *J. exp. Psychol.*, 38, 1948, 603-614.

[19] Hermann Aubert, Eine scheinbare bedeutende Drehung von Objekten bei Neigung des Kopfes nach rechts oder links, *Arch. Pathol. Anat. Phys. Klin Med.*, 20, 1861, 381.

TABLE V

FREQUENCY OF CHOICES OF TWO CRITICAL FIGURES IN EXPERIMENT VII:
HEAD TILTED 90° WITH UPRIGHT DETERMINED BY GRAVITY

Condition	3 A		3 B		3 C		All	
	U	RN	U	RN	U	RN	U	RN
I. Head right	8	1	7	2	6	3	21	6
II. Head left	5	4	6	1	7	2	18	7
Total	13	5	13	3	13	5	39	13
Percentage	72	28	81	19	72	28	75	25

In fact, no apparent increase in difficulty of recognition is at all observable when one's head and the object (for example, text) are both inverted, although this means 180° environmental disorientation. Yet all our experimental results have been running counter to observations of this kind. It was decided to try out our procedure under comparable conditions..

Procedure. O stands facing away from the scene and bends his head and trunk over until his head is upside down. In this position he looks through the aperture at the scene. To prevent extreme discomfort he is allowed to lift his head up after each test card is shown. In other respects the procedure is identical with the previous experiments in this series. One of the critical figures is now upside down and the other remains upright. In this experiment it was not necessary to have two conditions involving head tilt to different sides since the head was tilted 180°. About 30 Os were used.

Results. The results are given in Table VI. It can be seen that now there are about an equal number of choices of the two critical figures in every case. Chi-square does not reveal a significant departure from a 50:50 distribution of results.

This outcome is evidence for the role of both environmental and retinal factors, and in this case they are of about equal weight. This result is surprising. We expected a greater show of strength here of the retinal determinant. It is true, however, that in this experiment it has a greater effect than in Experiment V involving a 90° difference. In comparison with 29% in Experiment V, here the selection of the retinally normal critical figures approaches 50%.

TABLE VI

FREQUENCY OF CHOICES OF TWO CRITICAL FIGURES IN EXPERIMENT VIII.
HEAD TILTED 180°

	3 A		3 B		3 C		All	
	U	RN	U	RN	U	RN	U	RN
Totals	14	13	14	12	15	12	43	37
Percentage	52	48	54	46	55	45	54	46

DISCUSSION

As noted earlier, no attempt has been made to analyze the results for the different figures in terms of their special properties. It is obvious, however, that the different figures do not always yield the same results. Certain of the results for individual figures are not difficult to explain. Thus, for example, in the series using nonsense figures, in Experiment III, scene right, the version of Fig. 3C which remains retinally normal is such that the straight line portion of the figure is almost parallel to the upright of the scene. This coincidence seems to emphasize its difference, no doubt because now, within the scene, it is a 'straight' and no longer a 'tilted' figure. Only 2 of 14 Os select this version. Similarly in the scene-left variation this same line of Fig. 3C is now almost horizontal within the scene, with the same effect. Here only 7 of 21 Os select this version. In both cases the ratio in favor of the environmentally upright figure is greater than the ratio for the experiment as a whole. In general, however, we have not made a thorough analysis of specific figural changes which occur as a result of changes of orientation.

In all experiments the results have been given in terms of frequency of reaction to one or another alternative or critical figure. Since in no case is one alternative selected 100% of the time, the question arises whether we may assume that there are personality or other individual factors of some kind which so interact with the determinants in question that the percentages reflect individual differences? If this were the case the distribution of preferences should be such that some Os choose one alternative in all cards whereas others consistently choose the other. This, however, is not the case. There are few cases of such consistency. Consequently it would seem that the results must be explained by the fact that in some experiments the opposition of determinants is so balanced that the choice between the alternatives becomes more or less a chance sampling (Experiment VIII for example). In others, where one determinant *is* more effective, it is probable that the advantage is not so great that it can override chance factors completely.

PHANTOMS IN PATIENTS WITH LEPROSY AND IN ELDERLY DIGITAL AMPUTEES

By MARIANNE L. SIMMEL, Illinois State Psychopathic Institute and College of Medicine, University of Illinois

When a person has lost an arm or leg in an accident, or as the result of surgery, he typically continues to experience the lost extremity for some time. It was Mitchell who first coined the term *phantom limb* for this enduring experience of the limb which is no longer present.[1] On first awakening from the anesthesia such a patient may not believe that the leg actually has been removed until he can convince himself by looking under the covers. Even when he knows beyond doubt that it is gone, the foot of the amputated leg may itch and he may reach down to scratch it. He may perceive the bed sheets on it; experience a mild, perhaps pleasant tingling within it (a phenomenon which Henderson and Smyth regard as basic),[2] and even, though rarely, sense pain. He may 'feel' that he can wiggle his fingers or toes, flex or extend his wrist or ankle, and that he can perform these movements more or less at will. Despite his knowledge of the amputation which has been performed, the patient may 'forget' and reach out with his missing hand to grasp something, or to steady himself, or he may step on the phantom foot and fall.

Phantoms are not limited to limbs. Removal of other body parts—such as fingers, toes, the nose, eyes, nipples, and genitals—results in similar experiences. Nor is amputation proper a necessary condition for the appearance of phantoms. They appear typically after complete denervation by plexus-lesion or spinal-cord transection and, more rarely, in certain types of peripheral nerve section as well as occasionally after cerebral lesions. Under these conditions, the affected body part is of course still present, but the neural pathways which ordinarily mediate the sensations reported have been eliminated. Thus, a paraplegic patient may experience sensations of tingling in his legs; he may 'feel' that he can move the toes of the

---

[1] S. Weir Mitchell, Phantom limbs, *Lippincott's Mag. pop. Lit. Sci.*, 8, 1871, 563-569.

[2] W. R. Henderson and G. E. Smyth, Phantom limbs, *J. Neurol. Neurosurg. & Psychiat.*, 11, 1948, 88-112.

paralyzed extremities at will; or he may 'feel' the limbs drawn up either over his abdomen or to one side while in fact they are outstretched.

It should be emphasized that such experiences are normal and practically universal sequelae of the conditions indicated above. With time, the phantoms undergo changes which have been described elsewhere,[2] but they do not disappear altogether for a long time, often for many years, and they may endure throughout life in a very considerable proportion of patients. Mitchell estimated that 95% of all amputees experience phantom limbs, and this figure has been revised in an upward direction by many authors since his time. Everyone of Bors's 50 paraplegic patients examined from 2 mo. to 8 yr. after injury reported phantom experiences.[4] Painful phantoms are clinically recognized as the almost invariable result of plexus-lesions, although there has been no systematic study involving a large number of patients so affected. The number of phantoms in patients with peripheral nerve transection or cerebral lesions is small, however, probably because the requisite type and locus of lesion occurs rarely; furthermore, the patients with cerebral lesions are difficult to examine, and other symptoms are much in the foreground.

In general, phantoms also occur when a previously denervated structure is subsequently amputated, though the reported facts are somewhat less clear for these conditions and may at first sight appear contradictory. In paraplegic patients in whom amputation of one leg later becomes necessary because of gangrene, the amputation-phantom seems to resemble the phantom of the contralateral paralyzed leg, although the details of the published material leave much to be desired. There are other patients who suffer such severe phantom pain after denervation, that the surgeon finally removes the affected limb in the hope of giving the patient relief—a vain hope, since the phantom pain characteristically persists unabated. It is of course possible that in these cases the pain itself is an important factor in the persistence of the phantom. A patient has recently been seen whose elbow was shattered by machine-gun fire, with resulting total anesthesia and paralysis below the elbow, and contraction of the fingers. The patient had a great deal of phantom pain and, several years after the original injury, a surgeon removed the useless arm, hoping to free the patient of his pain, but he continues to have a vivid, very painful phantom arm and hand. Many patients with painful phantoms after plexus-lesions have had identical experiences.

There are, however, some reports in the literature which seem to indicate that phantoms do not necessarily follow after amputation of a previously denervated structure. There is, in the first place, a case reported by Head of a patient with an ulnar paralysis and total loss of sensitivity in the little finger for 7 mo. before surgical removal of the arm.[5] The phantom hand which appeared after the amputation had only four fingers, with the little finger missing. In the light of later observations, including those to be reported below, Head's case would seem to constitute an exception rather than the rule. More important seems an observation of Pick

[2] M. L. Simmel, On phantom limbs, A.M.A. Arch. Neurol. & Psychiat., 75, 1956, 637-647.

[4] E. Bors, Phantom limbs of patients with spinal cord injury, A.M.A. Arch. Neurol. & Psychiat., 66, 1951, 610-631. Cf. also David Katz, Psychologische Erfahrungen an Amputierten, Bericht ü. d. VII Kongress f. exper. Psychol., 1922, 49-75.

[5] Henry Head, Aphasia and Kindred Disorders of Speech, 1, 1926, 489.

that phantoms were absent after amputation following frostbite in soldiers of World War I.[6] Unfortunately, systematic studies are lacking. To date the writer has examined two patients with amputation of digits following frostbite, only one of whom reported phantoms. Since the different time-element makes this a rather crucial question, a systematic investigation of frostbite-patients both with immediate and delayed amputation is required.[7]

There is one group of patients whose potential contribution to the phantom-limb problem has apparently not been recognized heretofore. These are patients who have lost fingers and toes by absorption due to leprosy. The clinical course of this process is gradual and altogether painless. At the outset there may be some paresthesias, such as tingling, numbness, or pins-and-needles sensations, but these are soon superseded by total anesthesia, beginning at the tips and spreading proximally. Then follows gradual absorption of bones and soft tissues, a process which extends painlessly over a number of years until the barest remnants of stumps are left—and even these may gradually disappear. The outstanding features of these cases are the following: (a) total sensory loss precedes loss of structure; (b) both types of loss are gradual, extending over years, although for a given structure the sensory loss takes much less time than the digital absorption; (c) both types of loss are altogether painless.

While the digits are painlessly absorbed, the anesthesia continues to spread proximally, up the arms and legs and sometimes involving the forehead and upper part of the face. When these patients injure themselves, as happens not infrequently, they experience no pain if the injury and its direct sequelae do not transcend the current limits of the anesthesia; but they do have pain if more than the current area of sensory loss is involved. In any case, these wounds heal only very slowly, if at all, and frequently become infected. The infection may spread, the whole limb may become edematous, and eventually gangrene may set in, necessitating amputation. The conditions of amputation and absorption for these patients thus differ in three important respects. First, some of the amputees have experienced a good deal of pain prior to amputation, although ordinarily not immediately before the surgery; in fact, by the time it is necessary to remove the limb the sensory level has usually advanced well beyond the level of amputation, to the extent that in many cases removal of the limb can be carried out without surgical anesthesia. Secondly, the loss of the extremity itself is sudden, even though gradual loss of digits and gradual sensory deprivation have usually preceded the amputation. Finally, absorption is limited to the digits, while the amputation usually causes greater loss.

Experiment I

The problem of the present study was to discover whether phantoms result as a consequence of absorption of digits in patients with leprosy. This condition differs from other conditions known to produce phantoms in the

---

[6] A. Pick, Zur Pathologie des Bewusstseins vom eigenen Körper: Ein Beitrag aus der Kriegsmedizin, *Neurol. Centralbl.*, 34, 1915, 257-265.

[7] For a somewhat different emphasis of these phantoms, see S. Feldman, Phantom limbs, this JOURNAL, 53, 1940, 590-592.

TABLE II

Results for Patients with Amputations and Absorbed Digits
(No Surgery on Absorbed Digits)

S	Amputation				Prior pain	Phantom	Absorbed digits	
	site	year	age	reason			site, etc.	phantom
6	right below knee. Patient is psychotic.	1930	24	absorption, gangrene secondary to leprosy	no info.	Vivid phantom following amputation. "Forgot," stepped on phantom and fell. Could wiggle phantom toes, but "got over that."	All fingers, both hands probably since long before 1950.	Apparently none; but says that "sometimes he misses little finger."
7	left below knee	1949	49	absorption of toes, ulceration secondary to leprosy	yes	Vivid. After amputation kept feeling for foot which was not there. "Forgot," stepped on it and fell. Could wiggle big toe, other toes felt turned under, foot turned. This corresponds to position before amputation. Prosthesis feels like "real leg," not wooden.	All fingers, both hands at least since 1930, perhaps longer.	None.
8	right below knee	1936	36	necrosis, tarsal bones, gangrene secondary to leprosy	yes	Vivid phantom, persisting, still can wiggle toes, still "forgets" and falls.	All fingers, both hands.	None; "all dead long ago."
9	left thigh	1953	61?	absorption, ulcer, osteomyelitis, secondary to leprosy	yes	Vivid phantom including toes, foot, calf; on several occasions when taking off shoes has taken shoe off prosthesis. Phantom coincides with prosthesis.	All fingers, both hands. Many years.	None; "I lost them once and for all."
10	left ankle	1955	56	absorption of toes; ulceration secondary to leprosy	none	Vivid phantom following amputation—foot felt as if there. Now intermittent, does not feel toes, cannot wiggle toes.	All fingers, both hands. Probably since about 1935.	None.
11	left below knee	1949	35	ulcers, infection, gangrene, secondary to leprosy	none	Vivid phantoms for both legs since amputation. Once "forgot" and almost fell. Sometimes when stump itches feels little toe itching too. No movement of toes.	Absorption of toes of both feet, especially big toes for many years prior to amputation of legs (since age 16?)	None of toes prior to amputation of legs. Now occasional itching of little toes, but only at times when amputation-stump itches also.
	right below knee	1949	35					

TABLE III

PATIENTS WITH AMPUTATIONS AND ABSORBED DIGITS (SURGERY ON ABSORBED DIGITS)

	Amputation						Absorption with surgery	
s	site	year	age	reason	Prior pain	Phantom	Site, surgery, etc.	Phantom
12	left ankle	1952	29	Trophic ulcers, osteomyelitis, secondary to leprosy.	some	Vivid. Voluntary wiggling of toes and raising of instep. Occasional burning sensation in arch of instep. Prosthesis "acts as if I have a real foot."	Absorption of fingers at least since 1938. Between 1945–53 remnants of all 10 fingers removed surgically after injury, infections, etc.	Vivid phantoms of all fingers (with possible exceptions of little fingers both hands). Feels occasional burning and itching at finger tips; still tries to pick up objects with absent fingers, especially fingers with right hand.
13	right below knee; left below knee	1953	51	Osteomyelitis secondary to leprosy.	yes	Vivid since amputation, did not know it had been performed until she looked. Some pain in foot, no telescoping. Sometimes wakes up at night and tries to get out of bed, only then remembers that there is no foot. Has fallen repeatedly.	Gradual partial absorption of all fingers between 1917–1947, repeated surgery on fingers of both hands.	Vivid phantoms. "I was going to pick up something with my fingers and I look and I say, 'Oh my God, there's no fingers.' I know I haven't got them but that's what happens sometimes."
14	right below knee	1953	52	Chronic perforating ulcers secondary to leprosy. (Diagnosed 1915 as psychotic but no marked disturbance or difficulty during interview.)	yes	Vivid, sometimes wiggles toes even though these had been amputated 10 mo. earlier; forgets amputation and falls. Prosthesis feels like real foot, patient can move foot in shoe.	Absorption of all fingers since 1935 with multiple surgery; thinks she had at least 4 or 5 operations on each hand.	Vivid phantoms, still reaches for things, feels finger tips, little and ring fingers most vividly.
15	Left below knee	1949	62?	Ulcers secondary to leprosy. (Psychiatric consultation (1952) with diagnosis of beginning senile deterioration, but no difficulty during interview other than defective speech due to missing nose.)	No info.	Vivid at first, kept trying to scratch foot; less vivid now, but still can wiggle toes despite their having been amputated in 1940 after partial absorption and infection. No telescoping.	Amputation, left toes prior to 1929 after partial absorption. Amputation, nose 1933 after partial absorption. Fingers of right hand slowly absorbed since about 1930. Amputation of right thumb 1940, ring finger 1952.	Phantom toes as part of leg left amputation phantom. Vivid phantom of nose, sometimes itches and patient still tries to scratch it. Vivid phantom for thumb, tries to feel things. Repeatedly emphasizes that he does not feel the other fingers, only the thumb and 4th finger which were amputated.
16	Right foot partial; right below knee	1936, 1943	36, 43	Perforating ulcer secondary to leprosy. (Coöperation not very good, patient is in bed, quite ill and rather grouchy.)	yes	"For a couple of years you'd still think I had it [leg]." "Felt it at least once." Prosthesis felt like own leg, thought he could move ankle and toes of prosthesis	Both hands severely clawed and fingers partially absorbed since at least 1934. Amputation, left thumb 1938. Tendon transplant, right hand (fingers largely absorbed) 1940.	No phantom, not even of amputated thumb "because it had been dead so long." No phantom of fingers of right hand.
17	left F₁, F₂,	1939	46	Contractures, infection, secondary to leprosy. Communication difficult, speaks very little English.	no info.	No phantom,	All toes largely absorbed, at least since 1942. 3rd and 4th toes right amputated, 1951.	No phantom except for 3rd and 4th toes right. "Yes, toes feel as if still present."
18	All toes	1953	53	Infection, ulcers, contractures secondary to leprosy.	none	Vivid phantoms, can move toes and at night feels bed sheet over toes.	Fingers of both hands largely absorbed (muscle transplants) Minor surgery recently.	No phantom.

are not available for all patients in this group, but for those obtained the facts are clear; the phantom leg includes the toes even though they had been lost through absorption prior to surgery.

These facts are difficult to account for unless we hypothesize that the cognitive schema is not all of one piece, that (in a sense) it has memory, or that the alterations which it undergoes along with the gradual absorption of structure are nowhere near as stable or permanent as was the original 'normal' schema which the individual had attained prior to his crippling illness. Once the occasion for any kind of phantom arises at all, *i.e.* upon amputation, the resulting phantom corresponds to the original schema, *i.e.* to the whole extremity, the complete digit, not just the part which was recently amputated. There is one case of Riese's to be discussed below, which supports this assumption with a somewhat different, though related observation. Let us note also that there is apparently some limit on the memory of the schema, on the reversibility of its changes. The two cases (17 and 18) who underwent plastic surgery of the hand long after absorption of digits was complete did not report phantoms of the fingers. To be sure, this is too small a number from which to draw conclusions with any finality, but the observation is at least suggestive. Nor do these two cases indicate whether it is time itself which is responsible, or whether there is some specific stabilizing factor which happens to take place during that time.

There are several additional variables which were investigated in the present study—and one in the study reported below—which might have been thought to be of significance in the occurrence of phantoms. To summarize briefly, the presence of phantoms in our Ss is not related to age at amputation (the youngest S was 13 yr. old at the time he lost his leg), nor to the time which has elapsed since the loss (the longest was 35 yr.). Nor does pain prior to amputation play a significant role. Of the 21 instances of amputation, there were 10 reports of prior pain and 7 reports of no pain, while in 4 cases this information could not be obtained. Phantoms were reported, it will be remembered, in each of these instances.

Experiment II

There is, however, one other source of variation which might conceivably influence the occurrence of phantoms. It would not be altogether unreasonable to conclude from the data presented in Tables I, II, and III that phantoms follow predictably upon loss of the leg or the foot, while they have less than an even chance to result after loss of digits, no matter how this loss was incurred. Such a conclusion would be erroneous. It so happens

that only fingers, toes, and nose can disappear by a mere process of absorption due to leprosy, while loss of the foot or leg is invariably due to amputation.[12] Nevertheless, while the literature is unambiguous about phantoms resulting from amputation of extremities, there has been no careful study of phantoms after amputation of digits. For this reason I decided to study such a group.

Subjects. The Ss were 31 men, residents at a home for elderly (largely World War I) veterans, and included practically all digital amputees living there at the time of the examination.[13] These Ss were somewhat older than the patients with leprosy, ranging in age from 51 to 82 yr., with 24 Ss in the 60-69 yr. bracket. Age at amputation ranged from 6-62 yr., with 18 Ss having lost the affected digits at ages between 30 and 49 yr. In only three Ss had the loss occurred less than 10 yr. prior to the present study. With few exceptions, the amputations were the results of accidents, primarily industrial accidents, with traumatic amputation or with injury so severe as to make surgical amputation necessary. In approximately half the cases, the surgery was performed on the day of the accident, usually within hours, and the wounds healed without becoming infected. These 31 Ss can be divided into two groups: those with total amputations of some or all fingers or toes, and those with only partial amputation of the affected digits.

Procedure. The procedure was essentially the same as that followed with the leprosy patients. In individual interviews of about 30 min. in length each S was questioned about the events leading up to the amputation and about his subsequent experiences. Where phantoms were reported, E attempted to elicit a detailed account of the characteristics of the phantom and any changes it might have undergone in the course of time.

Results. The relevant details for the 17 Ss with total amputation of digits are presented in Table V. The results are quite clear: all except two Ss give definite evidence of having experienced phantoms since amputation, either continuously or intermittently. This is true irrespective of whether fingers or toes are affected, of whether the amputation was performed on the day of the accident or delayed (even though in the latter case infection tended to be an added factor), and of whether there was pain or not immediately prior to or after amputation.

A few comments about the last four Ss listed in Table V are in order. The history of S 114, who reports no phantom experience, is somewhat extraordinary. According to this S, his right second toe had to be amputated because of an infection. At the same time he requested that the left second toe also be removed "to make it even," and the surgeon evidently assented. Ss 115 and 116 have bilateral amputations as a result of frostbite; with similar periods of delay before amputation S 115 reports no phantoms while S 116 reports persistent phantoms. This only supports our view that a

TABLE V

RESULTS FOR Ss WITH TOTAL AMPUTATION OF DIGITS

			Amputation data		Injury-amputation delay	Pain		Phantom
S	age	years since	digits*	cause		prior	post	
101	60	<1	RF₁	accident	same day	yes	yes	yes
102	40	24	LF₁	accident	same day	yes	yes	yes
103	16	49	LF₁	accident	same day	no	yes	yes
104	56	13	RF₁₋₂	accident	same day	no	no	yes
105	38	24	LF₂₋₃	accident	same day	yes	no	yes
106	23	37	RF₃	accident	same day	no	no	yes
107	32	47	LF₂₋₃	accident	same day	no	no	yes
108	21	39	LF₄₋₃	accident	same day	no	no	yes†
109	26	50	LT₂	accident	same day	yes	yes	yes
110	54	11	LT₃	accident, gangrene	7–10 da.	yes	yes	yes
111	37	29	RF₄	accident, infection	10 da.	yes	no	yes
112	39	42	RF₃	puncture wound, blood poisoning	3–4 wk.	yes	yes	yes
113	48	34	RF₃	accident, blood poisoning	6–8 wk.	no	no	yes
114	25	39	RT₂, LT₂	right "infection"; left "to make it even"	? wk.	yes	yes	no†
115	46	16	all fingers, both hands	frostbite	same day	no	yes	
					2–3 mo.	yes	no	no†
116	57	11	all toes, both feet	frostbite	4 mo.	yes	yes?	yes†
117	34	35	LF₃	accident, nerve section contracture	4 yr.	no	no	yes

* F₁, thumb, F₂, index finger, etc.; T₁, big toe, T₂, 2nd toe, etc.; R, right, L, left. All Ss reported themselves to be right-handed.
† Doubtful about accuracy of S's report.

detailed study of frostbite cases is eminently called for. S 117 underwent amputation 4 yr. after the original accident which had crushed the ulnar part of his hand with sensory loss and gradual contraction of the little finger. Despite this he has a persistent amputation-phantom—one more case which is not in agreement with the usual conclusion drawn from the case of Head previously cited.

The details for the 14 Ss with partial amputations of digits are shown in Table VI. The results are only slightly less clear-cut than those from the Ss with total amputation of digits. Again the large majority of Ss do experience phantoms, but a few apparently do not. It does not seem to make any difference whether two joints were removed, or one, or only the finger tips, and whether or not there was pain prior to the amputation. Probably

the delay between accident and amputation is not a factor either, although this is less clear in the present group. Of those who report no phantom, only one S underwent amputation on the day of the accident, but there are also a number of Ss who experience phantoms after longer delays.

The last three cases in this table need a few additional words. One might think that where only the tip has been removed from a finger it would be impossible to question the patient meaningfully about phantoms or to obtain answers with any degree of certainty. Yet Ss 129 and 130, who suffered the loss of finger-tips only seemed to have no difficulty in understand-

TABLE VI

RESULTS FOR Ss WITH PARTIAL AMPUTATIONS OF DIGITS

S Case	Age	years since	digits*	amount	cause	Injury-amputation delay	Pain prior	post	Phantom
118	38	14	LF$_{1-3}$	2 joints	accident	same day	no	yes	yes
119	43	19	LF$_{1-3}$	2 joints	accident	same day	yes	yes	yes
120	34	35	RF$_2$	2 joints	accident	same day	no	yes	no
121	31	29	LF$_1$	2 joints	accident	same day	yes	yes	yes
122	45	15	LF$_{2-4}$	2 joints	accident	same day	no	yes	yes
123	37	26	RF$_3$	2 joints	accident	1 wk.	yes	yes	no
124	41	21	RF$_3$	2 joints	accident, infection	3 mo.	yes	no	yes
125	15	47	RF$_{2-4}$	1 joint	accident	same day	no	no	yes
126*	31	39	LF$_{2-4}$	1 joint	accident	same day	no	yes	yes
127	14	52	LF$_2$	1 joint	"Shot it off because of terrible itching;" blood poisoning	4 days	yes	no	no
128	31	37	LF$_1$	1 joint	injury, infection, blood poisoning	7 wk.	yes	yes	no
129	47	4	LF$_2$	tip only	accident	same day	yes	yes	yes
130	62	4	RF$_{1-4}$	tips only	accident	2-3 da.	yes	no	yes
131	6	55	RF$_3$	2 joints	accident, blood poisoning	2-3 mo.	yes	no	no

* S 126 reported himself to be left-handed; all others claimed to be right-handed.

ing the questions and in giving apparently reliable reports to the effect that they did indeed experience phantoms. S 129 is of additional interest because the accident which necessitated removal of only the tip paralyzed the remainder of the finger. This patient tells not only of an amputation-phantom, but also of a phantom of the paralyzed stump, reporting that he "forgets" that it is stiff and still occasionally tries to use it. Finally, S 131 should perhaps not have been included in the table at all, since his amputation occurred when he was only 6 yr. old, and the fact that he claims never to have experienced a phantom is not surprising in view of the reported absence of phantoms after amputation in early childhood.[14]

Excluding this last case we find that 15 out of 17 Ss with total amputations of digits and 9 out of 13 Ss with partial amputations, i.e. a total of 24 out of 30 Ss, gave definite evidence of amputation-phantoms. It may

[14] Pick, op. cit., 257-265; W. Riese and G. Bruck, Les membres fantômes chez l'enfant, Rev. neurol., 83, 1950, 221-222.

be noted also that there was considerable doubt about the reliability of the information provided by 9 Ss, and that this group included 5 out of 6 Ss reporting no phantoms. One of these cases (S 120) not only reported that he had never experienced a phantom, but practically denied the fact of the amputation. Two others (Ss 114 and 127) had lost digits as a result of at least a degree of self-mutilation. Still other Ss seemed to be quite disenchanted once they were informed that the results of the examination would not be instrumental in increasing their pensions and, as a consequence, seemed to coöperate in only the most superficial manner. The data may, therefore, be regarded as representing a conservative estimate of the incidence of phantoms after amputation of digits. The real incidence is likely to be somewhat higher, perhaps approaching that of phantoms after amputation of other parts of the body.

DISCUSSION

In the presentation of the results of these experiments, the concept of the cognitive schema was introduced, and it seems difficult to account for phantom phenomena without this concept. While many questions necessarily remain unanswered at present, there are some relevant observations in the literature which merit brief discussion. These observations concern the absence of phantom limbs in individuals with congenital aplasias, in young children, and in the feeble-minded.

Congenital aplasias. Throughout the literature we find frequent reiterations of a comment by Pick to the effect that congenital aplasias and amputations in early childhood do not result in phantoms.[14a] As far as I have been able to discover, Pick's statement was based on his general understanding of the problem rather than on any direct evidence obtained from the examination of such patients. That he knew of the two cases reported by Souques and Poisot is at least doubtful.[15] The evidence which has accumulated since Pick's time supports his view, even though it is far from complete. To my knowledge there has not been any systematic study of patients with congenital aplasias. In fact, besides the case mentioned by Souques and Poisot I know of only one other patient who has been studied with this question in mind— a 10-yr.-old girl with congenitally absent arms who apparently had never experienced any phantoms.[16]

Amputations in early childhood. In 1950 Riese and Bruck published a brief note on 24 child-amputees.[17] They could find no evidence of phantoms of limbs removed before age 6 or 7 yr. One of the cases is of special interest in connection with our

[14a] Pick, *op. cit.,* 257-265.

[15] A. Souques and Poisot, Origine périphérique des hallucinations des membres amputés, *Rev. neurol., 13,* 1905, 1112-1116.

[16] Personal communication from John Steelman.

[17] *Op. cit.,* 221-222.

finding of phantoms when amputation is superimposed on absorption. The leg of one boy was amputated at age 6 yr., apparently without resulting in any phantom experience. Three years later, at age 9 yr., the stump was revised with extraction of some bony fragments, and following this operation a phantom appeared. Unfortunately, further details are lacking. In a recent reanalysis of Cronholm's data on 122 amputees, I found no incidence of phantom in those patients (four in number) whose amputations had been carried out before age 5 yr.[18] Two patients with amputation at this age still had phantoms of the right leg 19 and 34 yr. later, respectively. Souques and Poisot have one case of a 16-yr.-old patient whose hands had been amputated at 2 mo. of age and who reported never to have experienced phantom hands or fingers.[19] Dr. . . . Steelman reports that one of his patients, a man 38 yr. old who had lost his right arm when he was 3 yr. old, could not recall ever having experienced a phantom.[20] Finally, one of the writer's own cases, who lost two joints of the right index finger at age 6 yr., claims never to have experienced a phantom, but the reliability of this man's report is in doubt because of his inability to recall any childhood experiences.

Do these cases support the assumption that amputation in young children does not produce phantoms? In our opinion they do not. They indicate only that amputations among preschool children do not result in phantoms which persist over many years, nor in phantoms which are recalled at a later age, but they leave open the possibility of immediate and relatively short-lived phantoms which are not remembered later. This possibility is strengthened by stories familiar to every newspaper reader. From time to time we come across a story of a youngster who, perhaps as a result of an accident, has lost a leg or a hand and who, a day or two after the surgery when the story is written, "has not yet been told" about the amputation. Such reports would seem to indicate the presence of a phantom, because otherwise the child would not have to be told; he would miss his extremity as soon as he wakes from the anesthetic. Clearly, this is a problem for future investigation. Meanwhile all we can conclude is that phantoms which persist, or are at least recalled on later questioning, do not develop if the amputation takes place before ages 5 or 6 yr. We might speculate further that perhaps phantoms do not occur at all in the very young, that they have a more or less fleeting existence in children who are a little older, and that they develop stability only at the age at which other perceptual and perhaps mnemonic functions reach some degree of stability—at the age (6-8 yr.) at which for instance, children are capable of the visual-spatial organization requisite for learning to read and write.

[18] Börje Cronholm, Phantom limbs in amputees, *Acta Psychiat., Kbh. Suppl.,* 72, 1951, 7-310.
[19] *Op. cit.,* 1112-1116.
[20] Personal communication.

Amputations among the feebleminded. Although no systematic studies have been made, it is occasionally stated that phantom limbs are not reported by feebleminded amputees.[21] Several considerations are pertinent in this connection. First, low-level defectives, who do not develop beyond the intellectual level of the normal 5- to 6-yr.-old child, may be disqualified on those very grounds. Secondly, those mental defectives who have developed somewhat beyond that level may still lack the analytical capacities to discriminate a phantom, especially one which they have lived with' for some time. Add to this the difficulty of communication with such Ss about such a patently unreal topic. Finally, in addition to their original handicap, adult mental defectives have a lifelong experience of being regarded as foolish, and many may find it far simpler to deny phantom experiences than to lay themselves open to further ridicule. The difficulties of such a study therefore should not be underestimated.

If a stable cognitive schema is a prerequisite for the appearance of phantoms, it is not difficult to see why they may not occur in the very young, who have not yet attained such an organization. Similarly, those lacking the intellectual endowment for its attainment may not have phantoms. The schema of individuals with congenital aplasia must of course lack the missing body-parts, and for that reason phantoms are absent in persons so affected. We have pointed out that systematic evidence is lacking for some of these conditions. Meanwhile, however, the present attempt at a formulation of the problem has highlighted the importance of some crucial details and thus might aid future studies.

SUMMARY

Eighteen patients with leprosy were interviewed about phantom-experiences resulting from loss of limbs and digits following amputation and absorption. All patients had undergone at least one amputation, and 13 of the Ss had, in addition, suffered absorption of fingers or toes. The data obtained indicate the following:

(1) Phantoms characteristically appear after amputation.

(2) They do not result from absorption of digits as long as surgery is not superimposed on the process of absorption.

(3) They appear when the remnants of partially absorbed digits are amputated, but they do not seem to appear following plastic surgery of the hand long after absorption of fingers is complete.

(4) The phantoms which result from amputation of partially absorbed digits do not consist merely of the remnants that were surgically removed; they are phantoms of the whole digits prior to any absorption.

---

[21] Papilion, Des interprétations délirantes et des hallucinations chez les amputés aliénés, Thesis, Lyon, 1905; abstracted in *Rev. neurol.*, 13, 1905, 915.

(5) Upon amputation of a leg, the toes of which had been absorbed prior to the amputation, the phantom experienced is complete, including the toes.

(6) A control study of 31 digital amputees showed that phantoms normally follow upon traumatic or surgical amputation of digits.

(7) The results of this study indicate that the following variables do not influence the appearance of the phantom: (a) age at the time of loss, provided amputations in early childhood are excluded; (b) the time elapsed since amputation; (c) the nature of the condition which makes amputation necessary; (d) the delay between the original injury or infection and the time of amputation; and (e) the presence of pain prior to or following amputation.

PAST EXPERIENCE AND PERCEPTION: MEMORY COLOR

By S. Carolyn Fisher, Chester Hull, and Paul Holtz,
University of California at Los Angeles

Can past experience, or expectations based upon this, affect the 'seen' color of a present stimulus-object? The experiment to be reported is concerned with what appears to be a striking difference between the findings of two recent studies in this area.[1] The authors of the earlier study, Bruner, Postman, and Rodriques, arrive at the general conclusion that apparent color of a perceived stimulus-object may be shifted in the direction of a strongly expected color provided the conditions are such as to dim or impoverish the stimulus-color, and to hamper the process of comparing the stimulus-color with a comparison-sample. This view is in line with that presented earlier by Duncker.[2] If, on the other hand, the color 'signal' is clear and sharp, and the perceptual conditions highly favorable to its being matched with a comparison-stimulus, the existence of an expectation has little or no effect upon the equation. Harper, the author of the more recent study, presents findings which he interpreted as indicating that even under the latter conditions the presence of an expectation has a marked effect upon perceived color. Certain features of these results led us to the belief that factors, other than experiential, might be affecting the data. The purpose of the present experiments is to investigate further the role of past experience in perception under good perceptual conditions, and to isolate or suggest additional variables, if any, which might influence the color-judgments or equations.

In each of the experiments mentioned above, the stimuli were forms cut out of colored paper or other material, which were to be matched under various conditions to colored comparison-stimuli. The forms in some cases represented common objects of specific color, this 'experienced' color being different from that of the material from which the stimulus-forms were made. Paired with each such color-meaningful stimulus-object was a 'neutral' form, i.e. one which had no associated color, cut from the same material and more or less carefully matched in size and shape to the meaningful stimulus. The hypothesis to be tested was that a stimulus-object nor-

mally of a specific color would be matched to a comparison-color which was closer to the experienced or expected color than would be the case for the object of identical material for which no such expectation existed, *i.e.* the color of the expectation-stimulus would be some sort of a compromise between the actual stimulus-color, and the expected color. Duncker, in his study with this method, found that naïve Ss who viewed stimuli in an illumination which greatly reduced their color did make such a compromise in matching to a rotating disk of variable components. Bruner, Postman, and Rodrigues, with stimuli more carefully equated in size and shape, confirmed Duncker's finding in their experiments with 'low information' stimulus-objects. Their experiments included a contrasting series in which a group of Ss (Group IV) viewed the stimulus-objects in good illumination and made their comparisons under relatively favorable conditions for clear seeing: stimulus (orange paper) and variable color wheel (red to yellow through orange) were viewed through a reduction tunnel at a distance of 150 cm. and with a separation of 10 cm. between their proximal borders. Under these conditions the means of the matchings showed no trends discernible to inspection toward differential matching of stimuli with red or yellow expectation (*e.g.*, lemon, tomato), and neutral stimuli. We ignore for the time the results of their analysis of variance.

Harper's paper gives a criticism of the procedure employed by Bruner, Postman, and Rodrigues, a changed apparatus, and some markedly different results. He thinks that the procedure of these authors admitted of the operation of response-generalization, hence that their results cannot be interpreted univocally in terms of perceptual modification.[3] In his own experiment Harper employed an apparatus in which figures cut from orange paper were suspended in the center of a 2-in. aperture through which S saw a surface (*i.e.* a background for the figure) which varied from yellow to an orangish red.[4] Harper hoped that at a certain point the stimulus-figure would melt into the background so completely that S would be unable to detect it. In such cases there would be no doubt that the figure and background were being perceived

---

[3] In this connection, Harper cites C. C. Pratt (The role of past experience in visual perception, *J. Psychol.* 30, 1950, 85-107). We concur with this criticism of the Brunner, Postman, and Rodrigues investigation. The use of stimuli of exceedingly poor informational character would reduce the perceptual component and increase the knowledge factor in the solution of his task. Moreover, in the light of Joynson's findings (The problem of size and distance, *Quart. J. exper. Psychol.* I, 1949, 119-136) it would seem that their 80° separation between stimulus-object and comparison-disk would carry S's performance wholly into the area of memory or knowledge, especially as the indefinite stimuli would have little potency for retention. Moreover, the fact that a group of Ss who received very specific information regarding the actual identity of the material of the several stimuli, and who should hence have had a highly specific expectation, nevertheless behaved in the same fashion as did the non-informed subjects would seem to cast serious doubt on the expectation-interpretation. Such a result would lend weight to the suspicion that psychophysical factors of some sort were operative. These comments apply, however, to the Bruner, Postman and Rodrigues procedure with very unfavorable conditions for perception. They have little, if any, bearing on the "Group IV" experiments with clear and favorable comparison-conditions and very small angular separation of stimulus-figure and comparison-disk. It is these latter experiments which are most comparable with Harper's and our own work.

[4] R. S. Harper and C. R. Oldroyd, An inexpensive differential color-mixer, this JOURNAL, 65, 1952, 614-616. See also P. T. Young, Differential color-mixers, *ibid,* 66, 1953, 312 f.

as the same, and not merely responded to as such. He was not able fully to realize this goal, because his illumination of the stimulus-figures was not wholly uniform and S could not consequently, make the whole figure blend with any one background. When S reported this difficulty he was instructed to make the upper central part of the stimulus blend, and to ignore the rest of the figure.

Harper's stimulus-figures were similar to those used by Bruner, Postman, and Rodrigues in their experiments of Group IV with clear, easy conditions for perceptual comparison. They consisted of three pairs, all cut out of the same orange paper, each pair composed of a normally red object (an apple, a heart and a lobster) and a figure having no characteristic color, but having the same area and the same general contour-characteristics as its meaningful mate (an oval, an isosceles triangle, and a letter 'Y' respectively). Five Ss participated. They were instructed to signal when the figure (designated as "a *red* heart," an "oval," etc. merged with the background. Each S made two determinations for every figure. The tabulated results (expressed in terms of the degrees of red in the matching background) showed a strong trend for the normally red member of each pair to be matched to a redder background than was the case with the neutral member, both for naïve and for psychologically sophisticated Ss. In the total of 20 comparisons each for the 3 pairs, this trend was reversed only once. These findings obviously contrast strikingly with those of Bruner, Postman, and Rodrigues for their (Group IV) data with favorable comparison-conditions.

Certain features of the data, however, appeared to us to indicate that factors other than past experience were playing a considerable part in the readings. For one thing, the three non-meaningful figures vary greatly in their redness matchings. Again, if the matchings for the 'Y' are compared with those for the heart, the former are redder in 7 out of the 10 comparisons, and they average redder. The average and range for each of the figures, in descending order of redness of matched background, are: (1) Lobster, 156°, 132-170; (2) Apple, 141°, 121-161; (3) 'Y,' 111°, 66-135; (4) Heart, 105°, 75-130; (5) Triangle 60°, 15-111; and (6) Oval, 39°, 6-81. There are marked gradients between the apple and the 'Y,' and between the heart and the triangle. The difference between the results of Harper, and those of Bruner, Postman, and Rodrigues, suggest that a number of variables inherent in the apparatus may be effective, *e.g.* the size of the field in which the stimulus is viewed: the restricted, aperture-colored surround in Harper's arrangement as compared with the large area of stimulus and comparison-disk of Bruner *et al.*

PROBLEMS

Our experiments were undertaken to test further the factor of experience with relation to perception, and to determine the possible effect of other variables. They include three parts: (I) A repetition of the work of Harper, with some improvements in apparatus; (II) An investigation of the effect of experience, using stimulus-objects which varied only in orientation; and (III) An investigation of a series of 'meaningless' forms, varying in area and in complexity.

Apparatus and procedure. Essentially the same apparatus as that employed by

Harper was used in all of the experiments. S was seated in front of a gray screen at a distance of 4 ft., kept constant by means of a chin rest. Mounted in the screen was a blackened tube 6 in. long and 2 in. in diameter, through which S saw the (varying) background and the small stimulus-figure suspended in front of it. Hence our experiments, like Harper's, presented a stimulus-figure seen in a small immediate color-surround which took on the appearance of a film- or aperture-color in the subdued illumination of the room.

The following modifications were made in Harper's apparatus to improve the equalization of the illumination of the stimulus-figure. A disk-shaped piece of solid plastic, 1 in. thick and 4 in. in diameter, was placed over the distal end of the viewing tube. Four small light bulbs were placed at equal distances in holes drilled into the rim of the disk, out of S's view. These lights threw an even illumination on the stimulus-figure, which was cemented flatly to a thin clear plastic disk. The cement used in Experiment I darkened the backs of the stimulus-materials which slightly reduced their brightness, hence the results of this experiment are not directly comparable with those of Experiments II and III. The disk bearing the stimulus-figures was held in place by three metal prongs which protruded backward from the periphery of the thick disk containing the light bulbs; these prongs fitted into holes in the margin of the stimulus-disk. The stimulus-figure was thus held flatly in the middle of the 2 in. colored background, produced by the reflection of the large variable disks, which was to be matched to it. The distance between the figure and the plane of the four light-bulbs was 1 in. In this way it was possible to get the stimulus to blend into the background when the latter approached it in hue and brightness, and no frosted plastic was needed to blur the figure. S usually found that at the matching point the stimulus-figure tended to disappear and a 2 in. field of uniform film color remained. To step up the saturation of the intermediate orange hues we illuminated the large disks, which were the source of the background, with a 100 w. bulb covered with an orange filter. We also introduced a variable A.C. transformer in the circuit feeding the four small bulbs to attain a level of stimulus-illumination which matched the intermediate color range of the background.

EXPERIMENT I

Experiment I. was essentially a repetition of Harper's work. The Ss were seven psychologically naïve volunteers from undergraduate classes in psychology.

Procedure. The stimulus-figures were representations of the objects Harper specified. All were cut from orange paper (Ostwald, hue 5 p.a.). Measurements of height and width of the stimulus-figures were approximately equal throughout the series. The areas were determined by projecting the figures onto 1/10-in. graph paper, counting the squares covered, and reducing this number by the magnification-factor. The following measurements were obtained: Heart: area, 0.749 sq. in.; perimeter, 3.58 in.; P/A, 4.78;[5] Triangle: 0.800, 4.26, and 5.3; Apple: 0.828, 3.725, and 4.5;

---

[5] P/A = ratio of perimeter to area. This value gives a rough measure of complexity. See J. E. Hochberg, Henry Gleitman, and P. D. Macbride, Visual threshold as a function of simplicity of form, *Amer. Psychologist*, 3, 1948, 341 f.

Oval: 0.745, 3.08, and 4.13; Lobster: 0.545, 4.57, and 8.39. 'Y': 0.502, 4.05, and 8.07, respectively.

In presenting the stimulus-figures, we randomized the direction of background-change, *i.e.* from red (*FR*) or from maximal yellow (*FY*). Every meaningful figure was alternated with every meaningless one, the whole series being run twice at one sitting: thus, every *S* made six *FR* and six *FY* matchings for each of the six figures. Means for the first and second runnings of the series showed no consistent differences; there was no evidence in the data of a practice or a fatigue effect.

Instructions. The instructions employed in Experiment I were as follows.

(Meaningful figures): This will be a red heart [apple, lobster]. It's really redder than it looks in this light. As I change the color of the background, at a certain point the figure will blend into it. Please say "now" when the figure merges into the background.[6]

(Non-meaningful figures): This will be a triangle [oval, 'Y']. As I change the color of the background, etc., as for meaningful figures [. . . .].

E then uncovered the observing tube, exposing the figure against a 'yellow' ground (180° *Y* + 180° *R*) or a red one (360° *R*), and changed the color of the ground at as fast a rate of speed as *S* could take without feeling insecure in his judgment; this helped in reducing after-image and contrast effects. At *S*'s signal, *E* replaced the gray cover on the tube, and took down the reading from the scale. He then set the disks for the opposite background and repeated the procedure. The stimulus-figure was then changed and the *FR* and *FY* readings made as before. *S* thus had from 10 to 25 sec. of rest between matchings; if he was aware of after-images the time was extended. A few trials with an extraneous figure were given *S* at the outset to familiarize him with his task. None of the *S*s suspected that the figures were actually the same in color.

Results. Table I gives the mean background matches for each of the

TABLE I

MEAN DEGREES OF RED IN THE BACKGROUNDS MATCHED TO EACH FIGURE
(EACH MEAN BASED ON 84 OBSERVATIONS.)

	Heart-Triangle	Figure Apple-Oval	Lobster-'Y'
Meaningful member	284.85	290.57	294.75
Neutral member	288.50	286.52	290.75

figures. To evaluate the significance of the apparent reversal of the heart-triangle pair, and also that of the between-figures differences independent of the pairings, the data (summed scale readings) were subjected to a triple classificatory analysis of variance: figures × orders × *S*s, where order refers to the direction from which the background was changed (*i.e.* *FR* or *FY*). The results are reported in Table II.

---

[6] In this instruction, we deliberately emphasized the expectation-factor, planning to use less 'loaded' instructions in a later series with the aim of studying the possible effects of instructions on performance.

The term 'figures' should not be interpreted too literally since the variable of shape is confounded with the possible differential effect not only of meaning, but also of instructions. The significant 'S × order' interaction is interpreted to indicate that some Ss showed an 'anticipation' common in the method of limits others were consistently 'tardy.'

We computed t-ratios between each set of means; these are reported in

TABLE II

Summary of Analysis of Variance

Source of variance	df.	Sums of squares	F	(num./denom.)
(1) Figures	5	9582	6.34	(1/5)*
(2) Orders	1	4846	3.496	(2/6)
(3) Ss	6	48755	5.86	(3/6)*
(4) F×O interaction	5	1.066	.745	(4/7
(5) F×S interaction	30	9058	1.056	(5/7)
(6) S×O interaction	6	8313	4.85	(6/7)*
(7) S×O×F interaction	30	8581		

* Significant beyond the 1% level.

Table III. The error-term was derived from the figure × S interaction mean square, as suggested by Lindquist.[7]

Confining our attention to extra-pair differences (because of the above-mentioned confounding), the 'Y' is judged more red than either the oval or the heart, differences being significant. It is interesting to note that with the exception of the heart, the rank order of the magnitudes of red in

TABLE III

Showing "t"-Ratios Among the Various Figures With Regard to the Redness of Their Matched Backgrounds; Together with the Levels of Significance

	Lobster		'Y'		Apple		Triangle		Oval	
	t	signif.	t	signif.	t	signif.	t	signif.	t	signif.
'Y'	2.955	<.01								
Apple	3.003	<.01	.14	>.10						
Triangle	4.478	<.01	1.614	>.10	1.474	>.10				
Oval	5.91	<.01	3.04	<.01	2.906	<.01	1.431	>.10		
Heart	7.106	<.01	4.24	<.01	4.101	<.01	2.626	<.02	1.194	>.10

the backgrounds which matched the figures is the same as that obtained by Harper. Study of the direction of the extra-pair differences suggests that neither area nor the P/A measure taken alone will account for the results. Inspection of within-pairs figures reveals that, of each pair, the figure matched to the redder background was the larger of the pair both in area and in the P/A measure. This holds for the reversed heart-triangle means.

[7] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 164 f.

EXPERIMENT II

The purpose of Experiment II was to assess the role of experience in perception under the same general experimental conditions as before, but employing more Ss and using stimulus-figures with which the variables other than expectation (meaning) between the experimental and the control-figures were minimized.

Procedure. In this experiment we used as 'meaningful' figure the symbol of the Red Cross—the height and breadth of the figure being 1 in. with arms $\frac{1}{4}$ in. wide. The 'neutral' figure was this identical cross rotated through an angle of $45°$ so that it appeared as a sign of multiplication. (X). Thus except for orientation, the meaningful and the neutral figures were identical.[8]

The cross was cut from the same orange paper as had been used in the previous experiment. A few trial runs with two Ss indicated that the identity of the two figures might be detected, hence we inserted two additional figures (the apple and the oval from the previous experiment) in the sequence, taking two measurements of each, but not retaining these for the statistical analysis.

Ten orders were drawn up which randomized the sequence of presentations of the cross (CR) and the (X) and the direction of background-change $(FR$ vs. $FY)$. Each S made five judgments under each of the four possible conditions (*i.e.* CR-FR; CR-FY; X-FR; and X-FY). The score assigned to each S for statistical analysis was the sum of the scale readings obtained under each of the four conditions. The deception-figures were presented alternately in the sequence as follows: one was always given first, and at about the fifth, tenth, and fifteenth trial in the series of 20 trials. The introduction of the extraneous figures (apple and oval) was successful in preventing S from suspecting that the cross and the X were identical figures—at least none of them reported or remarked that the two were the same.

The Ss, 37 in number were undergraduate students in psychology who were required to serve two hours as experimental subjects during the semester. The procedure was the same as in Experiment I, except for slight changes in the instructions and the asking of four questions after the termination of the 20 matchings.

Instructions. The instructions were as follows.

We are doing some experiments on color contrast. We are using four figures, an apple, a circle, a Red Cross symbol, and a letter X, and you will be asked to match them to a background. (Then, for the first few showings): This will be an X [cross, red apple, oval]. I am going to change the color of the background, and

---

[8] It is impossible to be sure, at present, that the change in orientation can be assumed to imply no changes beyond itself. Köhler and others have attributed an anisotopic character to visual space. The factors which determine the horizontal-vertical illusion may be effective here. P. G. Cheatham (Visual perceptual latency as a function of stimulus brightness and contour shape, *J. exp. Psychol.* 43, 1952, 369-380) found a highly significant difference in the latency of a square lying on its side, and the same square turned through $45°$ to form a diamond. On the other hand, Bitterman et al. (M. E. Bitterman, John Krauskopf. and J. E. Hochberg, Threshold for visual form: a diffusion model, this JOURNAL 67, 1954, 205-219) found no significant form-threshold differences between an horizontal and rotated square, and between an upright and rotated cross (such as we used).

at a certain stage it will match the color of the figure. The figure may seem to blend or merge into the background at this point. Please say 'now' when the background is the same color as the figure. [After the first few showings only two first and last sentences were used. Whenever the cross was to be shown, the following was substituted for the first sentence]: This will be the Red Cross, the well-known symbol of the Red Cross Society. It's really redder than it looks in this light.[*]

The four questions asked the Ss at the conclusion of the trials were: (1) What do you think the purpose of this experiment was? (2) Will you rank the figures in the experiment from the most red to the least red? (3) Does that mean that the cross (or the X) looked redder to you as you saw it in this experiment? (4) Would you be willing to bet that you matched the cross (or the X, if ranked redder) to a redder background than you did the X (or cross)? Question (2) was asked of all the Ss; the other three questions were asked only of the last 28 Ss serving in the experiments.

Results. Table IV gives the mean degrees of red in the backgrounds which were matched to the cross and the X, for both directions of change,

TABLE IV

MEAN DEGREES OF RED IN BACKGROUNDS MATCHED TO THE CROSS
AND THE X FOR EACH DIRECTION OF CHANGE

Order	Cross	X
FR	222.71	222.21
FY	235.23	233.79
Total	230.18	228.00

and totals. The smaller amount of red in the matching backgrounds as compared with data of Experiment I is attributable to the use of a different cement. The cement employed in Experiment I darkened the figures.

A triple classificatory analysis of variance was performed, the cell entries being the summed scale-readings for each S. Table V contains the summary. An examination of the data yielded, as the most likely basis for the S × figure (expectancy, or meaning) interaction, a tendency on the part of a majority of the Ss to match the cross to a redder (or to a less red), background, as compared with the X, in rather a consistent fashion. The extent of this tendency was as follows: 16 Ss matched the cross redder in both series (FR and FY); 5, matched the cross redder in the FR series only, 5 in the FY; and 11, matched the cross redder in neither series (*i.e.* matched X redder). On the assumption that if only random factors are operating the population for each of the four cells should have been the

[9] We anticipated the possibility that S might find these instructions ambiguous. The instructions might mean, "Say 'now' when the background is the same color the Red Cross really is," or "when the background is the same color the Red Cross appears to be under present conditions." We planned, if this question arose, to instruct S to use the latter criterion. None of the Ss, however, raised the question.

same, we calculated the chi-square, throwing out one of the cases of the first group to avoid fractions although thereby slightly lowering the result. The chi-square (3 df.) was 7.8, significant at the 5% level.

In an effort to interpret the above distribution, we examined the results of the questions to determine whether the matchings reflected a belief or hypothesis relating to the color of the cross as compared with the X. If some Ss had accepted the suggestion in the wording of the instructions whereas others had reacted negatively to it, this might conceivably have accounted for the scatter. The following facts appeared: (1) No S guessed the real aim of the experiment. (2) Of the 16 Ss who match the cross redder in both series (FR and FY), 7 rank it as redder, 3 rank it the same,

TABLE V

SUMMARY OF ANALYSIS OF VARIANCE FOR THE CROSS AND THE X

Source of variance	df.	MS	F	Level of confidence
(1) Figures (expectation)	1	10.81	1.85 (1/6)	.20
(2) Orders (FR, FY)	1	1685.25	43.07 (2/5)	.001
(3) S (S)	36	54.45		
(4) FXO	1	2.49	.93 (4/7)	
(5) SXO	36	38.50	14.31 (5/7)	.001
(6) SXF	36	5.85	2.17 (6/7)	2.5
(7) SXFXO	36	2.69		

6 rank it less red. (3) Of the 20 Ss who rank the cross redder (the remainder being uncertain), 7 match it redder in both series, 5 in one series, and 8 match it less red in both series. (4) Of the 6 Ss who would, or "probably would," bet that they had matched the cross redder, 1 matched it redder in both series, 2 in one series, and 3 in neither series. (5) Of the 6 Ss who would bet they had matched the X redder, 4 matched the cross redder in both series, 1 in one series, while none actually matched the X redder.

Thus, although possibly a majority of the Ss reacted with or against the instructions on the conscious level, it is abundantly clear that the Ss' rankings did not predict their matchings, nor did their matchings predict either the certainty or the direction of their rankings. There is the possibility that the Ss used different criteria in making their matches, e.g. some equating in terms of chroma, others of brightness; but it is difficult to imagine why this should have affected the cross and the X differentially. Nor is it easy to think how the different orientation of the X as compared with the cross could have affected some Ss and not others. In conclusion, we are wholly at a loss to explain the bimodality of the distribution of the Ss. We cannot, of course, exclude the possibility that for some Ss, the

factor of expectation influenced the perception of redness in the cross. If this actually happened, expectation functioned without the Ss' being aware of it, *i.e.* expressing it in their rankings or other communications. Indeed, it must often have functioned when S expressed a contrary attitude or expectation.

EXPERIMENT III

In Experiment III the stimuli-objects were four meaningless figures, cut from the same orange paper as before. Thus they carried no suggestion of any particular color. It was hoped that they would serve to generate further hypotheses concerning the psychophysical variables on the basis of the color equations obtained. Ease of measurement was also a consideration in their selection.

Procedure. Two of the figures were disks of diameters 1 and 0.81 in. respectively; they will be referred to as large disk (LD) and small disk (SD). The other two figures varied mainly in complexity. One of them was a 1 in. disk with four square excisions cut from the periphery at equal distances; each cut-out was 0.25 in. long per straight side, one side, of course, being part of the curved circumference. Thus the contour was made angular, and also much lengthened. The fourth figure consisted of two disks 0.56 in. in diameter, slightly overlapped so that the lateral extent was 1 in. The latter two figures will be referred to as the notched disk (ND) and the overlapped disk (OD) respectively. The dimensions are given in Table VI, along with the results.

The method was that of average error; the Ss themselves varied the background by

TABLE VI

MEASUREMENTS OF STIMULUS-FIGURES TOGETHER WITH THE DEGREES OF RED IN THE BACKGROUNDS MATCHED TO EACH

	LD	ND	SD	OD
Area (sq. in.)	.785	.553	.518	.468
Perimeter (in.)	3.1416	5.4	2.55	2.79
P/A	4.002	10.05	4.93	5.96
Mean degrees red (no. obs'ns, 150)	263.39	252.11	247.12	239.92

pulling strings which were attached to the variable color-disks. The strings were made long, and the length to be pulled was varied in random fashion to prevent the use of positional cues. The instructions were the same as those for the X in the preceding experiment. Every stimulus-figure was matched 10 times, 5 FR and 5 FY as regards direction of background-change. The 10 matches for each figure were obtained successively, direction of background-change being randomized. The order of the four figures was randomized among the Ss. Fifteen Ss were used, similar to those of Experiment II in manner of selection and level of sophistication.

Results. Table VI gives mean findings, together with measurements of the stimulus-figures. We ran a double classification analysis of variance

using the summed scale-readings for each S. (Removing the order effects would have been possible; this would have increased statistical significances by decreasing the size of the error-term.) Table VII gives the results of the analysis.

Using the error-term derived from the interaction-term,[10] t-ratios were determined for differences among the means for the four figures. Table

TABLE VII

Summary of Analysis of Variance

Source of variance	df.	MS	F
1. Figures	3	450.36	49.59 (1/3)*
2. Ss	14	48.343	5.32 (2/3)*
3. Interaction F×S	42	9.0811	

VIII gives the differences and the ts. All of the latter are significantly beyond the 1% level of confidence.

These results confirm the findings of Experiment I in that significant differences are obtained between figures which are neutral as regards expectation, or meaning. The only systematic figure-difference which parallels the difference in magnitude of red in matching background, is that of area. Failure to find a clearcut increase as a function of the increased com-

TABLE VIII

Mean Differences, with t-Ratios for the Four Figures (df. = 42)

		LD	ND	SD
ND	Dif.	6.27		
	t	5.698		
SD	Dif.	9.04	3.23	
	t	8.216	2.936	
OD	Dif.	13.04	6.77	4.00
	t	11.851	6.153	3.635

plexity (P/A measure) may be due to a selection of figures which do not show a sufficient range of complexity or which are too much confounded with the measure of area. It is interesting to note that significant differences were obtained between figures with relatively small differences in area.

DISCUSSION

It would appear from our results that a number of psychophysical variables affected the levels of brightness, saturation, and the hue at which our test-figures merged into the limited area of the background employed.

---

[10] Lindquist, *op. cit.,* 164 f.

With regard to these variables, we were able to do little more than confirm their importance; we did not succeed in segregating any of them. Whether in addition experiential variables played a part in producing changes on the attributive level was not established. We were actually involved in problems of the effect of form and contour on difference thresholds in which color and saturation, as well as brightness, were implicated. To some extent these problems are similar to those which have recently been investigated by Bitterman *et al.*[11] These experimenters have maintained that the threshold for the perception of form is considerably affected by area, and by complexity. This latter factor is related to P/A, but especially to another variable designated as 'critical detail.' They have not as yet been able to find a satisfactory method of segregating and evaluating critical detail, but they identify and test certain examples of it: *i.e.* interior angles, with length of arms as magnitude of critical detail. They state that angular size, with especial reference to extremely acute and obtuse angles, is an important aspect of the problem of critical detail.[12] It seems possible that factors which affect the visibility threshold might also have some effect on difference thresholds at higher levels of background illumination. Since in all cases the stimulus-figures which have been used in experiments which aim to determine the effects of experience by the use of some sort of matching techniques have such differences, it is important to assess these figural variables before results can safely be attributed to experience.

Another variable which needs investigation in connection with such methods as Harper's and our own is size of aperture. A 2-in. aperture viewed from 4 ft. is well within the range of sizes for which the Fry-Bartley findings with respect to border interference apply. A sharp border such as that enclosing our area might have a considerable effect on borders of smaller enclosed figures. Thus, Bartley remarks: "The threshold for a disk on a slightly darker ground decreases as the area of the disk is increased" (up to about 4° of visual angle).[13] Inasmuch as brightness changes were present in the variation of our background, the conditions cited by Bartley could conceivably have had something to do with our findings relating to the effect of area (Experiment III). Fusion at lower

---

[11] Bitterman, Krauskopf, and Hochberg, *op. cit.*, this JOURNAL 67, 1954, 205-219; John Krauskopf, R. A. Duryea, and M. E. Bitterman, Threshold for visual form: Further experiments, *ibid.*, 427-440.

[12] It is interesting to note that Werner, in his studies of contour, concluded that angular parts of contours have more "psychophysical energy" than circular parts, and that they resist extinction (Studies on contour: I. Qualitative analysis, this JOURNAL 47, 1935, 40-64, esp. 43, 46, 48.)

[13] Howard Bartley, *Vision, a Study of its Basis*, 1941, 9.

brightness-levels would of course also mean greater redness in the matching mixture.

We are at a loss to account for the differences between Harper's results and our own. It is barely possible that the frosted plastic used by Harper to make his stimulus-figures blend with the background may have dulled the colors a trifle and favored 'memory color.' There is also the possibility that though his pairs of figures were equal in area (he does not state his method of equating them), the upper central parts which S frequently used as the basis of his judgments were not. For example, the heart and the triangle could have been distorted in such a way as to leave the area larger for the heart. It is not unreasonable to conjecture that the order of complexity of Harper's figures might have corresponded roughly to the order of redness or darkness of their matching backgrounds. Moreover, there is the difficult variable of 'salient detail' which might fairly include angular irregularities in contour, such as the corners of the triangles vs. the roundness of the heart; the bulges, projecting stem, etc. of the apple, as compared with the regularity of the oval; the pointed and cleft claws of the lobster vs. the more regular outline of the 'Y.' It is clear that Harper's order of matching for his figures does not follow any one of these variables in a simple and direct way, but it is also obvious that something other than past experience is varying among, and perhaps even between, his three pairs of stimulus-objects. Whether these factors are adequate to account for the large differences between his meaningful and meaningless figures is by no means certain.

In spite of the differences in viewing conditions and background-area, our data are closely similar to those obtained by Bruner, Postman, and Rodrigues for their Group IV (composed of Ss who worked under optimal viewing and comparing conditions) in that our means, like their's, contain no intelligible trend with regard to meaningful vs. neutral figure-colors. We were interested in noting that when the results of Group IV were subjected to an analysis of variance, both meaning and shape of stimulus-figure—even within such narrow limits as ovaloid to ellipsoid—became significant sources. A statement of shape X meaning interaction would have lent added interest. In view of the possibility that color matching is sensitive to slight differences in stimulus-pattern, it is not out of the question that the lack of discernable trends in the means arose from a complex combination of such psychophysical variables. Until these variables are effectively isolated, their role cannot properly be determined and the possibility of their contribution to differences in color equations attributed to past experience should not be underestimated even for rela-

tively poor conditions for stimulus-discrimination. Finally, this line of research may throw further light on the problem posed by Pratt and Harper, of distinguishing between perceptual modification and response-generalization in assessing the role of past experience in the perceptual area. If it is possible to distinguish experimentally between the two concepts in terms of the influence of psychophysical variables as opposed to learning and motivational factors, the term *perception* may gain back some of the specificity of meaning it has lost in recent years.

SUMMARY

Three experiments are reported which investigate the influence of certain psychophysical variables vs. the effect of past experience (or expectation) on the perception of the color of stimulus-figures physically identical in color. The experimental situation was modeled upon that of Harper. In all experiments, Ss viewed an orange colored figure at a distance of 4 ft., suspended in the center of a 2-in uniform field or background of color which was varied between red and (reddish) yellow through the intermediate oranges, one of which was a close match to the (constant) stimulus-color. Ss task was to signal when the background color matched the color of the stimulus-figure.

Experiment I was a repetition of Harper's work, with a few improvements in apparatus. Three pairs of stimulus-figures were used. One of each pair was a cut-out representation of a normally red object, while the other was a neutral form of size and shape similar to its meaningful partner. We failed to confirm Harper's findings that Ss in all cases matched the meaningful figure to a much redder background than they did the color-neutral one. The differences obtained by us were very small, and with one pair of figures they showed a reversal, the neutral figure being matched redder. Psychophysical factors in the stimulus-figures appeared to be the major sources of variation.

In Experiment II two stimulus-figures were used. These were identical in size and shape, the meaningful one being the well-known Red Cross of the Red Cross Society, the neutral one being the same rotated through 45°. While the cross was matched to a slightly redder background, the differences were very small and hardly significant. The Ss were questioned concerning their impressions, if any, of the relative redness of the figures; while some of the Ss matched the cross redder and others showed the opposite tendency, their matchings did not predict their beliefs (if any) regarding the relative redness of the stimuli, and vice versa.

In Experiment III none of the stimulus-objects was 'meaningful;' the

psychophysical variables of area and complexity of contour were studied, though we did not succeed in securing a systematic variation of any factor. Significant differences in matching redness appeared among the figures; area being the most consistently related to degree of redness with larger areas 'judged' redder. Further work with systematic variation of the stimulus-objects is required for adequate interpretation of the differences.

The findings are discussed in relation to the work of Harper, and that of Bruner, Postman and Rodrigues, being in fairly close agreement with the most comparable data of the latter investigators. Possible relationships to recent work of Bitterman and his co-workers are suggested.

THE EFFECT OF AUDITORY AND VISUAL BACKGROUND
ON APPARENT DURATION

By I. J. Hirsh, R. C. Bilger, and B. H. Deatherage,
Central Institute for the Deaf, St. Louis, Missouri

A general theory of perception requires the identification of appropriate dimensions for patterning in the several sense modalities. Visual patterns are formed from changes or gradients in brightness and color as a function of one or more of the three dimensions of space. It is thus that we speak of the physical referents for vision as *objects*. Auditory patterns, on the other hand, arise from changes in loudness and pitch as a function of time and it is more meaningful, therefore, to label the physical referents for auditory perceptions as *events*. These ideas are not new, but they need emphasis. The three dimensions of space can be and have been controlled and discussed fairly readily, but time seems more difficult to handle. As one approach to time as a perceptual dimension, this paper concentrates on the study of time as duration and not as a background-dimension in which patterns are generated.

The psychophysics of time-perception has been concerned primarily with the apparent duration of filled and empty intervals on the one hand, and with rhythmic patterns on the other. Most of the classical experimental work has been reviewed by Boring[1] and by Woodrow.[2] Stimuli used in the studies on duration have been presented through several sense-modalities.

We have been impressed, however, with the intimate relation between time and hearing that is suggested by two kinds of evidence. The first concerns the obvious dependence of auditory perception on patterns that are generated in time, as suggested above.[3] Secondly, adults who have heard normally and then have lost their hearing complain not only about the loss of communication but also report that the world about them appears to be quite dead and that life seems to have lost its on-going character.[4] These two considerations suggest the hypothesis that one's concept of time may depend in large part on auditory stimulation.

Although sounds can be placed in space by hearing alone, such localizations are in tactual or visual space. Contrariwise, visual and tactual phenomena may change as a function of time and give rise to such complex forms as rhythm and movement, which may have an auditory basis.

Problem. The present study tests the hypothesis that perceived time varies with the level of auditory stimulation. Does the apparent duration of a stimulus depend in some way on the background against which it is perceived? Specifically, we wished to know whether the auditory background, against which a tone or light is presented, affects the duration of the reproduced interval of these stimuli. If auditory input were capable of effecting changes in apparent duration, then there remains the further question whether input to another sensory modality has a similar effect. Visual stimulation was chosen to test this second question.

Method. The procedure and apparatus were so arranged that the observer (*O*) could reproduce the duration of a stimulus (tone or light) that was presented to him. The ambient conditions of light and noise or dark and quiet were so controlled that they could be held constant or be varied independently during the periods of *O*'s stimulation and of his response. Under the 'control conditions,' the ambient light and noise in the experimental room were the same during the periods of stimulation and response. Under the 'experimental conditions,' the ambient backgrounds during the periods of response differed from those during the periods of stimulation.

Apparatus. Fig. 1 shows a block diagram of the apparatus used. The 'start-relay' was energized manually by *E* and was deënergized after a pre-set number of pulses had passed through the counter. *O*'s 'response-button' energized the 're-sponse-relay' which again turned on the stimulus hence a stimulus was also present during the entire period of response. This response-button also operated an electric timer, allowing the duration of *O*'s response to be measured to 0.1 sec.

A 250~ tone at a sound-pressure level (above 0.0002 microbar) of 80 db. was used as the auditory stimulus throughout these experiments. Ambient noise was obtained by mixing the output of a white-noise generator with that of the tone-generator and presenting both through the same loudspeaker.

The visual stimulus was provided by a Sylvania R-1131C flash tube with an input of 90 v. The tube was fed by an amplifier that was normally biased to cutoff and was placed in operation by a countervoltage from a battery that was applied whenever the 'start-relay' or 'response-relay' was energized and the 'light-tone selector' was set on 'light.' The illuminated end of the flash tube was mounted behind a piece of sanded plexiglass that covered a hole of 5/8-in. diameter, cut in the center of a large, gray plywood panel. This board rested upright on the loudspeaker cabinet and the *O* faced both from his seated position. *O* sat 10 ft. from the loudspeaker and light-board; the light was on a line horizontal with his eyes and the axis of the loudspeaker was 2 ft. below this line. The ambient light was provided by a 100-w. photoflood lamp suspended from the ceiling.

The observations were made in an anechoic chamber which provided not only

a controllable sound-field without reflections, but also a convenient source of darkness.

Procedure. O was seated in a chair on the suspended wire-mesh floor of the anechoic chamber. He was told that after the word "tone" or "light" was spoken through the intercommunication system, the stimulus named would be presented for a certain length of time. When it disappeared, he was instructed to press and to hold his button down (which turned on the same stimulus) for the same length of time. He was further instructed not to count or to beat a rhythmic pat-
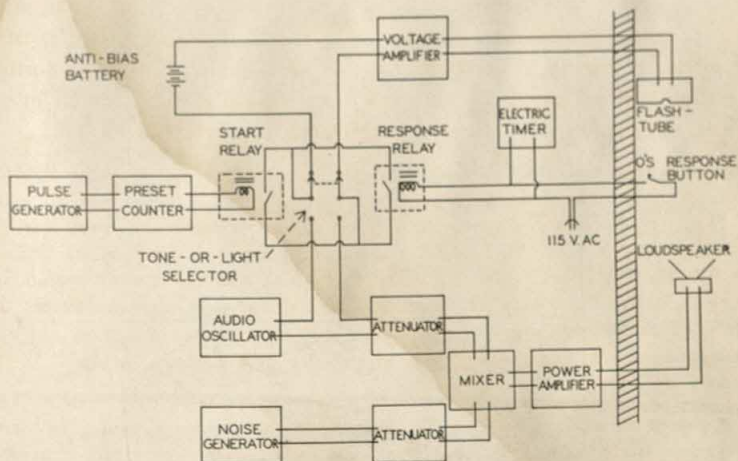
FIG. 1. BLOCK DIAGRAM OF APPARATUS

tern, but only to judge the overall duration as a unitary event. Each experimental session consisted of a random series of four presentations at each of five different durations (1, 2, 4, 8 and 16 sec.) under a given set of ambient conditions. Following a rest, in which O left the chamber and walked around for a few minutes, a second session was run. Each session lasted between 10 and 15 min. and no more than two such sessions were completed in any morning or afternoon.

Four 'control' conditions were used for each of the two stimuli, tone and light. In these, the ambient conditions were the same during periods of both stimulation and response; indeed they remained the same throughout the single session. These four were: light and quiet (LQ), light and noise (LN), dark and quiet (DQ), and dark and noise (DN).

The first set of 'experimental' conditions was characterized by constant ambient light or darkness during both stimulation and response, but with a change in the auditory ambient conditions from quiet during stimulation to noise during the response, or vice versa. These conditions obtained for both stimuli, tone and light.

In the second set of experimental conditions, the ambient noise or quiet was held constant but the ambient light changed from dark during stimulation to

light during the response, or vice versa. Other variations were introduced follow-
ing preliminary analysis of the data from these first three sets of conditions.
These variations will be described after discussion of the results of the first
three experiments.

Results. The results of the experiments under control conditions are
given in Fig. 2. The duration of the stimulus is shown along the abscissa
and the duration of the response is shown on the ordinate, as the percentage
of the duration of stimulus. If all responses were completely accurate,
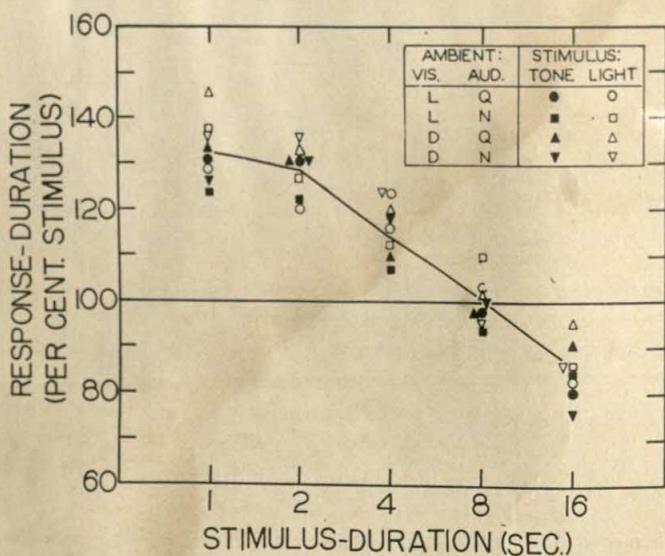
FIG. 2. RESPONSE-DURATION VS. STIMULUS-DURATION
Every point is the mean of four responses made by each of 8 Os.

duplicating the stimulus-durations exactly, the points would all fall
along the horizontal line at 100%. The four different shapes of symbols
indicate the four different combinations of ambient conditions, which re-
main unchanged through the periods of stimulation and response. The
filled symbols show results for the tonal stimulus, while the open symbols
refer to the light. Each of the symbols represents a mean based on four
individual responses made by each of eight Os.

It can be seen that durations of 1, 2, and 4 sec. are on the average
overestimated, that a duration of 16 sec. is underestimated and that 8 sec.
would have been called in the classical literature an "indifference-interval."
Among the classical experiments, those of Woodrow are similar to ours,

but he did not report a universal tendency for the underestimation of long intervals or the overestimation of short intervals.[5] In a later study, he reported an average indifference-interval of about 0.6 sec., and concluded, moreover, that the indifference-interval was independent of the range of stimuli because he used different groups of Os for the different durations.[6] Our indifference-interval, shown in Fig. 2, appears to be at least one whole order of magnitude greater and we doubt that the difference is to be accounted for as a difference of filled versus unfilled intervals. Woodrow has suggested, however, that in work on the absolute impression of a duration, the indifference-interval may very well be in the middle of whatever range of durations is used.[7] We need not dwell on this lack of agreement between our results and those of other investigators; the importance of the present study has to do with differences between the results in our control and experimental conditions. The experimental conditions do not differ from the control conditions in psychophysical procedure or range of stimuli.

A comparison of those control conditions represented by circles with those represented by squares, or of triangles with reversed triangles, shows no clear separation between responses made under conditions of quiet and those of noise. A similar comparison of conditions represented by circles or squares with those represented by either of the two types of triangles shows no difference between responses made in ambient light and those made in the dark. If either visual or auditory stimulation were capable of affecting the perceived time, we could not see such effects clearly here because any changes brought about would obtain equally well during stimulation and response. Accordingly, we arranged the first of a series of experimental conditions in which the auditory and visual ambient background was different during the response from what it was during stimulation.

In Fig. 3, we see the effects of changing the ambient auditory conditions as we move from the period of stimulation to the period of response for both tones and lights. While the acoustic change was made, the visual environment (either light or dark) remained the same during stimulation and response. The filled symbols show the means of responses made during 'noise' after stimuli had been presented during 'quiet,' while the open symbols show the reverse; namely, responses made during 'quiet' after stimuli had been presented during 'noise.' The separation between filled and unfilled symbols is obvious and clear, and furthermore applies equally

[5] Woodrow, The reproduction of temporal intervals, *J. exp. Psychol.*, 13, 1930, 473-499.
[6] Woodrow, The temporal indifference interval determined by the method of mean error, *ibid.*, 17, 1934, 167-188.
[7] Woodrow, Personal communication, March 11, 1955.

well for tones and lights and for ambient conditions of light or darkness.

The curve labeled "control" is not an average of the data represented by the two experimental curves, but rather is the curve of Fig. 2, inserted here to facilitate comparison of the results of these experimental conditions with the average of responses made when the ambient conditions were the same for periods of stimulation and response. It appears that a response made in the presence of noise to a tone or light that was pre-
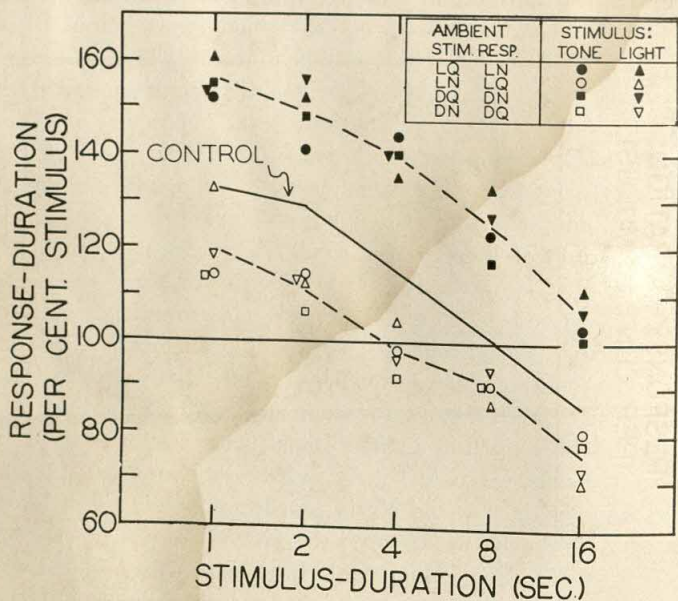
FIG. 3. RESPONSE-DURATION VS. STIMULUS-DURATION
Filled symbols show the results when stimuli were presented during quiet and responses made during noise; open symbols show the results under reverse conditions.

sented in the quiet is longer than a response made in the quiet to stimuli presented in noise. These comparisons hold not only between the two experimental conditions, but also between either one and those control conditions in which either noise or quiet obtained for both stimulation and response.

Having established that a change in the acoustic background against which a stimulus is presented will result in a change in its apparent duration, we proceeded to a similar experiment to ascertain whether changes in the visual background would produce similar changes in apparent duration. Fig. 4 shows the results of this experiment. The curve drawn is again

the control curve of Fig. 2 and is not intended to show the central tendency among the points of Fig. 4. Here we see that changing the visual ambient conditions from dark to light or from light to dark as we move from stimulation to response does not yield results that are significantly different from those obtained under the control conditions. The filled symbols refer to those conditions in which the change was made from dark to light, while the open symbols refer to the reverse conditions, changing from light to dark. A statistical analysis (see below) revealed that the data ob-
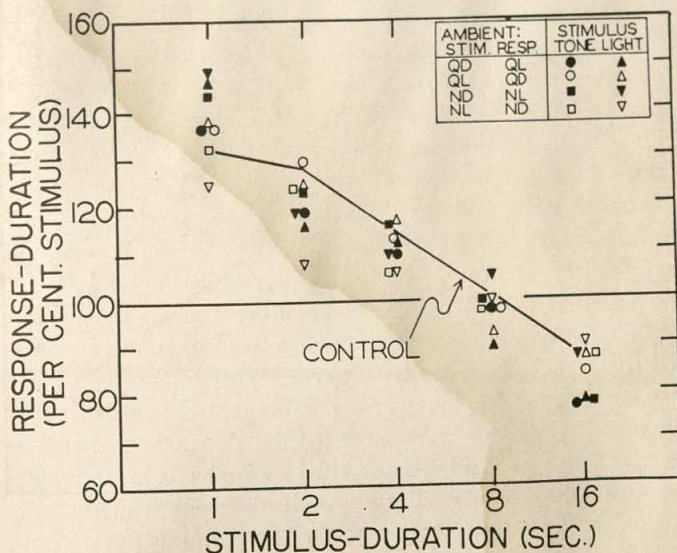
FIG. 4. RESPONSE-DURATION VS. STIMULUS-DURATION

Filled symbols show the results when stimuli were presented during dark and responses made during light; open symbols show the results under reverse conditions.

tained under these conditions of change in ambient light are not significantly different from those obtained under the control conditions. Furthermore, no difference was found between responses made in the dark following stimuli presented in the light (open symbols) and those made in the light following stimuli presented in the dark (filled symbols). Thus, it appears that under the conditions of the present experiments changes in the acoustic background effect changes in the apparent duration of stimuli, while changes in the visual background do not.

It is important to note that the data in Figs. 3 and 4 are representative of both tones (circles and squares) and lights (triangles). Thus, we are

not studying merely auditory duration but perhaps a perceptual dimension that transcends the particular sense modality being used. Psychological time appears to transpire at about the same rate in the dark and in the light, but it appears to run at quite different rates during noise (90 db.) and during quiet. The data of Fig. 3 suggest that time runs slower in the quiet than in the noise. A convenient way to express and remember the direction of this finding is to state that the quiet is dull and boring, while the noise provides a more interesting background.

Even in an anechoic chamber, a quiet condition does not mean no sound. The over all ambient noise level was probably less than 30db., but it is difficult for us to measure this on our sound-level meters. When the external level is very low, O is much more aware of his own internal sounds: the pulsing of blood, rumblings in the stomach, etc. If these differences between 90 db. and the limiting condition called "quiet" are of such magnitude as shown in Fig. 3, then there must be a relation between the speeding-up effect of the noise and the difference in noise-levels. Another series of experiments was initiated, therefore, in which the quiet condition was replaced by noise-levels of 72, 78 and 84 db., while the noise-level corresponding to 'noise' remained at 90 db. The experiments were carried out in exactly the same way, with the exceptions that no dark conditions were used and the stimulus was always the tone. Separate functions like those of Fig. 3 reveal a large amount of variability, but the pooled results for all Os over all durations show a systematic decrease in the difference between responses made in the noise after a stimulus is given in the quiet and those made in the quiet after a stimulus is given in the noise, as the noisiness of the 'quiet' approaches that of the 'noise.' These results are summarized in Fig. 5. The difference between responses made against a 90-db. background and those made in the quiet, shown at 'Q,' is the average difference between all filled and unfilled circles in Fig. 3, that is, the difference between quiet-to-noise and noise-to-quiet conditions that obtain only for tones presented and responded to in ambient light. Thus, the data show not only that the duration of a response in noise following stimulation in quiet is greater than the duration of a response in quiet following stimulation in noise, but also that this difference decreases as the difference between the noise-levels decreases.

Analysis of results. Initially these experiments were done with four trained Os, who were aware of the structure of the experiments and of the time-intervals being used. To verify the results obtained on this group,

the entire series of experiments was repeated with a group of four naïve
Os. The statistical treatment of the data was directed towards answering
two questions: Do Figs. 2–5 represent the data adequately? Are the re-
sults shown comparable for both groups?

The analysis of variance used on each set of data was essentially the
same, a repeated-measurements analysis in which the error terms were
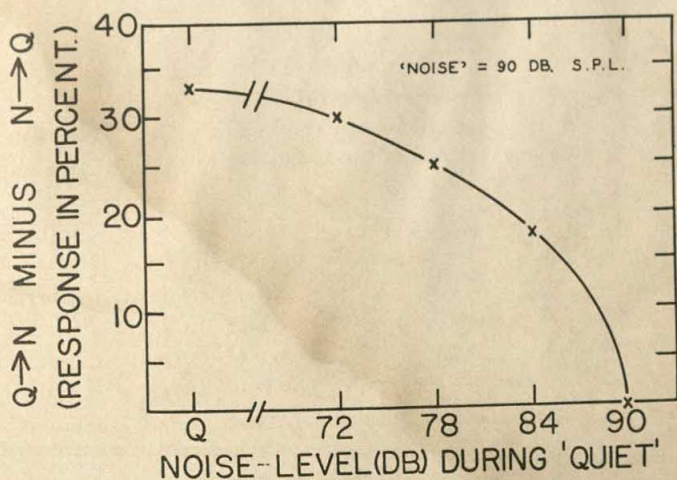
FIG. 5. DIFFERENCE BETWEEN RESPONSE-DURATIONS MADE DURING NOISE AFTER
TONE WAS PRESENTED IN QUIET AND DURING QUIET AFTER
TONE WAS PRESENTED IN NOISE
Every point represents an average of all trials for all Os at all durations.

pooled estimates from the two groups.[8] These analyses are summarized in
Tables I and II.

A difference between times was found in all four analyses. This result
is related to the tendency of Os to overestimate the 1, 2, and 4-sec. intervals
and to underestimate the 16-sec. interval.

In the three experiments in which both lights and tones were the stimuli
presented, a difference between the means of responses to the light and
tone was found only in the control conditions. A separate analysis of the
two groups of Os indicates that this difference occurs only in the naïve
group, accounting for the interaction of stimuli and groups. Since the
difference between stimuli was found only for the naïve Os under control

[8] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and
Education*, 1953, 156-189.

conditions, the conclusion that the Os responded differently to lights than to tones does not seem warranted.

The fact that no difference was found between the control conditions supports the results discussed above; namely, that no differences between conditions will appear when O reproduces the stimulus in the same environment in which it was presented.

In the experiments in which the ambient noise was changed between presentation of the stimulus and its reproduction by O, a difference was

TABLE I

ANALYSIS OF VARIANCE

Summary of control and experimental data (Figs. 2, 3, 4).

Source	d.f.	Control conditions		Q-N vs. N-Q conditions		L-D vs. D-L conditions	
		mean sq.	F	mean sq.	F	mean sq.	F
Groups (G)	1	6.4553	—	4.1633	1.00	1.2190	—
Os within G (O)	6	8.0842		4.1486		6.5582	
Conditions (C)	3	.7648	1.19	54.5613	15.72*	.99489	2.28
CXG	3	.4846	—	1.2342	—	.3492	—
CXO	18	.6418		3.4710		.4164	
Stimuli (S)	1	2.5794	25.85*	1.7287	2.10	.0034	—
SXG	1	1.6288	16.32*	.0320		.0346	—
SXO	6	.0998		.8221		.8314	
CXS	3	.4319	2.48	.1032		.2360	—
CXSXG	3	.6166	3.54	.1260		.0907	—
CXSXO	18	.1744		.7300		.3802	
Times (T)	4	39.7342	27.29*	35.4889	23.12*	44.2617	22.18*
TXG	4	.9494	—	1.2415	—	.5419	
TXO	24	1.4560		1.5348		1.9956	
CXT	12	.4147	1.66	.3582	—	.5766	1.76
CXTXG	12	.2133	—	.3153	—	.1008	—
CXTXO	72	.2496		.5979		.3281	
SXT	4	.2288	1.30	.3301	.160	.3532	1.06
SXTXG	4	.0749	—	.0342	—	.3864	1.16
SXTXO	24	.1753		.2057		.3335	
CXSXT	12	.3332	2.01	.4106	1.73	.2385	1.06
CXSXTXG	12	.2495	1.50	.2955	1.24	.2335	1.04
CXSXTXO	72	.1660		.2375		.2257	
Total	319						

* Significant beyond the 5% level.

found between the noise-to-quiet and quiet-to-noise conditions: the quiet-to-noise conditions resulting in significantly longer responses, as shown in Fig. 3. No difference was found, however, between those conditions in which the ambient light was altered between stimulation and reproduction. This analysis supports our finding that the acoustic environment used affects the apparent duration of both lights and tones, while the light environment used does not.

A summary of an analysis of variance on the experiments in which the level of 'quiet' was varied appears in Table II. The part of this analysis

critical for the results shown in Fig. 5 is the conditions-by-('quiet') level interaction (C × QL), which is significant. Further analysis of this C × QL interaction indicates that the difference between conditions becomes significantly smaller as the difference between the levels of quiet and of noise is reduced.[9]

The significant condition-by-time (C × T) interaction indicated in

TABLE II

ANALYSIS OF VARIANCE

Summary of experimental data (Fig. 5).

Source	df.	mean sq.	F
Groups (G)	1	8.0813	4.61
Os within G (O)	6	1.7545	
Conditions (C)	1	62.1591	60.15*
C×G	1	4.2987	4.16
C×O	6	1.0334	
Times (T)	4	11.6078	12.96*
T×G	4	3.1236	3.49*
T×O	24	.8958	
C×T	4	1.5551	4.58*
C×T×G	4	.9214	2.72
C×T×O	24	.3392	
Quiet level (QL)	2	3.2612	5.36*
QL×G	2	.1580	—
QL×O	12	.6084	
C×QL	2	1.9932	4.79*
C×QL×G	2	.8256	1.98
C×QL×O	12	.4162	
T×QL	8	.3694	1.70
T×QL×G	8	.2460	1.13
T×QL×O	48	.2171	
C×T×QL	8	.8173	2.56*
C×T×QL×G	8	.1993	—
C×T×QL×O	48	.3190	
Total	239		

* Significant beyond the 5% level.

Table II is attributable to the fact that the difference between conditions was most pronounced for the 4, 8 and 16-sec. intervals. Also, the significant condition-by-time-by-('quiet') level interaction seems related to the fact that the tendency for the longer time intervals to be affected most by the change in acoustic environment increases as the difference between quiet

[9] D. B. Duncan and R. G. Bonnor, Simultaneous confidence intervals derived from multiple range and multiple F-tests, Mimeographed report presented to American Statistical Association, Sept., 1954.

and noise decreases. The significant time-by-groups interaction is attributable to the differing responses of the groups to the 4-sec. interval.

Concerning the significant differences between 'quiet' levels, the 78-db. level is the deviant. This result cannot be explained in terms of the variables over which we had control.

Discussion. It is well known that the psychophysical study of time is extremely sensitive to procedural variations. The present studies have been limited to one procedure, involving the method of reproduction, without control of the interval between the cessation of the stimulus and the beginning of the response. Furthermore, we have employed a range of durations that differ from those used by most other investigators. Although we have not made a direct test, it would seem that disagreement between our general results, under control conditions, and those of previous investigators could be explained on the basis of differences in procedure.

The primary concern here, however, is with changes in an apparent duration that are brought about by changes in the environment of the *O.* Only two kinds of environmental or ambient stimulation were manipulated: the auditory and the visual. On the other hand, conditions like temperature, odor, and other kinds of sensory input, even though not controlled, may be assumed to have remained more or less constant throughout an experimental session. We assume that these environmental sources of stimulation constitute different kinds of background whose function, among others, is to pace the *O*'s psychological clock. Because of the fact that in this study only acoustic, and not visual, ambient stimulation seems capable of eliciting *marked* changes in the rate at which the psychological clock operates, we suggest that it is the auditory background that is used when available as a calibrating monitor for our clocks.

Particularly in those studies of time that have involved the perception of rhythm, the role of the motor system has been emphasized. Perhaps one of the shortcomings of the method of reproduction in this kind of a study is that it confounds the apparent duration of external stimuli, such as the tone or light, with the apparent duration of the muscular tension that accompanies the holding down of a button. The button had a very light spring and *O*s were instructed to "turn on the light or tone" with the button and attend to the duration of the stimulus that was thus turned on; but nonetheless several *O*s reported that it was button-holding more than the external stimulus that controlled the duration of their responses. Almost all *O*s agreed that this was particularly true for the light-stimulus.

There did not seem to be as intimate a connection between the button and the light as there was between the button and the tone. We hope to repeat these studies with a method of constant stimuli, that judgments of the durations of stimuli may be separated from durations in which the *O* is doing something.

The basic difference between visual and auditory background-stimulation with respect to changes in apparent duration, which constitutes the main finding of these studies, is restricted to these experimental conditions. We expect this conclusion to be general, but we must point out that the light- and noise-levels are not strictly comparable. If we utilize the comparisons between visual and auditory scales recently put forth by Stevens, we see that our noise of 90 db. does not strictly compare with the intensity of light since 90 db. is much higher than average conversation and the light in the experimental room was a little too weak for comfortable reading.[10] Also, the darkness *can* be thought of as having a much lower level than the quiet. A white bright homogeneous field might be more comparable to our environmental noise, and that extreme condition must be checked before we can feel secure about the generality of the findings. On the other hand, the data of Fig. 3 have been repeated with the noise-level at 80 db.

The problem concerning the kind of auditory or visual stimulation that might be used as a background has not been considered except in preliminary fashion. Similar results might be obtained if tones instead of noise were used, but it seems almost certain that the results would have changed considerably had a more definitely patterned auditory stimulation, such as music or speech, been used. Simple room-illumination appeared to be a reasonable analogue for the noise, but again had we used patterned events in visual stimulation, the results would also probably be different. The temporal and spatial characteristics of ambient stimulation will be the subject of future investigations.

SUMMARY

Eight *O*s were instructed to respond to a tone or light, whose duration was randomly set at 1, 2, 4, 8, or 16 sec., by pressing a button that again turned on the tone or light for a length of time considered by *O* equal to the original stimulus-duration. Control conditions under which such responses were made included the four combinations of light or dark with quiet or noise as background or ambient conditions. These visual and

---

[10] S. S. Stevens, Decibels of light and sound, *Physics Today*, 8, 1955, 12-17.

auditory environments remained the same during periods of stimulation and response. Under control conditions, durations of 1, 2, and 4 sec. were overestimated, 16 sec. was underestimated, and 8 sec. was reproduced quite accurately.

Experimental conditions were so arranged that the dark or light environment during stimulation could be changed to light or dark during the response, or such that quiet or noise could be changed to noise or quiet. Under these conditions, darkness or light had no effect on the apparent durations of tones or lights. Stimuli presented in the quiet, however, elicited responses made in the noise that were much longer than those responses made in the quiet following a stimulus presented in the noise. These very clear differences decreased as the difference between the noise-levels in the two periods was decreased.

The dependence of apparent duration on the level of auditory stimulation and the lack of dependence of apparent duration on the level of visual ambient stimulation is presented as evidence for a strong relation between apparent or psychological time and level of auditory stimulation.

PERCEPTION OF OVERLAPPING AND EMBEDDED
FIGURES BY CHILDREN OF DIFFERENT AGES

By Lila Ghent, New York University

This investigation is concerned with the child's perception of figures that are not clearly set apart from each other. Interest in this problem stems from reports that such perceptual tasks present special difficulty to brain-injured patients and to young, normal children.

Since the early work of Poppelreuter and of Gelb and Goldstein, various types of overlapping or embedded figures have been used in attempts to define perceptual effects of brain injury.[1] Goldstein has suggested that many of these effects may be subsumed under a concept of difficulty with the articulation of figure and ground, manifested in all aspects of the patient's behavior.[2] On the basis of Goldstein's hypothesis, Werner and Strauss used overlapping and embedded figures to study the performance of mentally retarded, brain-injured children, and have found abnormal performance in this group.[3] More recently, such tasks have been used by Teuber et al. in work with brain-injured children and adults.[4]

Reports on perception in normal children suggest that young children also exhibit difficulty on perceptual tasks in which figures are not clearly set apart from each other. Osterrieth has concluded that children between 4 and 7 yr. of age perceive a complex geometric figure as a relatively undifferentiated whole dominated by some details.[5] Although Osterrieth has shown clearly that young children find a limited number of forms in a particular complex figure, it is difficult to generalize from this work since only one figure was used. Witkin found that normal children show great difficulty in discovering geometric forms embedded in a series of complex figures, modified after Gottschaldt.[6] Both studies indicate that the young child responds less

[1] Walter Poppelreuter, *Die psychischen Schädigung durch Kopfschuss im Kriege 1914-17*, 1918, 16 and 210; Adhémar Gelb and Kurt Goldstein, *Psychologische Analysen hirnpathologischer Fälle*, 1920, 1-561.

[2] Goldstein, *The Organism*, 1939, 131-155.

[3] Heinz Werner and A. A. Strauss, Pathology of figure-ground relation in the child, *J. abnorm. soc. Psychol.*, 36, 1941, 236-248.

[4] H. L. Teuber, Neuropsychology, in M. R. Harrower (ed.), *Recent Advances in Diagnostic Psychological Testing*, 1950, 30-52; H. L. Teuber, W. S. Battersby, and M. B. Bender, Performance of complex visual tasks after cerebral lesions, *J. nerv. ment. Dis.*, 114, 1951, 413-429. The overlapping, realistic figures shown in Fig. 1 are those developed at the Psychophysiological Laboratory.

[5] P. A. Osterrieth, Le test de copie d'une figure complexe, *Arch. Psychol. Genève*, 30, 1944, 205-353.

[6] H. A. Witkin, Individual differences in ease of perception of embedded figures, *J. Personal.*, 19, 1950, 1-15.

adequately than adults to tasks involving the differentiating of figures not clearly set apart from each other.

The purpose of the present investigation was two-fold: (a) to define the presence and extent of the difficulty of a complex perceptual task for normal children of different ages (Experiment I), and, if such a difficulty can be established; (b) to explore some characteristics of the stimulus-objects that may give rise to the difficulty (Experiment II). In previous work with normal and brain-injured subjects (Ss), the unitary and generalized character of the difficulty has been emphasized, with relatively little analysis of the characteristics of the task that may be related to poor performance. Knowledge of the stimulus-conditions eliciting poor performance in normal children may help to clarify the nature of the defect in the brain-injured groups.

EXPERIMENT I

Subjects. Ninety-nine children between the ages of 4 and 13 yr. were tested. There were an approximately equal number of boys and girls at each age; the total number of children tested in each of four age-groups is shown in Table I. Testing was done at a child care and community center situated in a moderately low income area in New York City.[7]

Procedure. The pictures used consisted of overlapping line drawings of objects presented in the order shown in Fig. 1. The pictures were placed face down on the table, and S took one picture at a time from the pile. Most Ss handled the pictures in such a way that they appeared in the orientation shown in Fig. 1.

The instructions were as follows:

Tell me what you see in the picture. If you don't know the name of one of the things, show it to me with your finger. You can turn the picture any way you want, and you can look at it as long as you want. Be sure to tell me everything you see.

If S omitted some of the items, or looked to E for some comment after making a response, E usually said, "Do you seen anything else? If you don't know the name, show me with your finger." When S pointed to an item, he was asked to trace the picture with his finger. E demonstrated on an item that had already been identified by S what was required.

Scoring. The errors consisted mainly of omission and misnaming; there were very few errors of mistracing. Since the errors of mistracing were difficult to interpret, the analysis was limited to errors of omission. An omission meant that the child did not refer to a figure either by name or by pointing.[8]

---

[7] The author thanks the staff of the Bronx River Child Care and Community Center for their coöperation. Permission to test at this Center was arranged through Mr. Irving Brodsky of the Jewish Association for Neighborhood Centers and Mrs. Merl Hubbard of the Department of Welfare of New York City.

[8] The stringed instrument behind a similar instrument on Card 3 was not scored, since this item was drawn incorrectly, *i.e.* the F-opening of the instrument was not shown.

Results. Table I shows the percentage of *S*s within each group who failed to report each number of figures. The most striking result was the consistent decline in number of omissions with increasing age; a chi-squared test showed this trend with age to be significant at the 1/10% level.
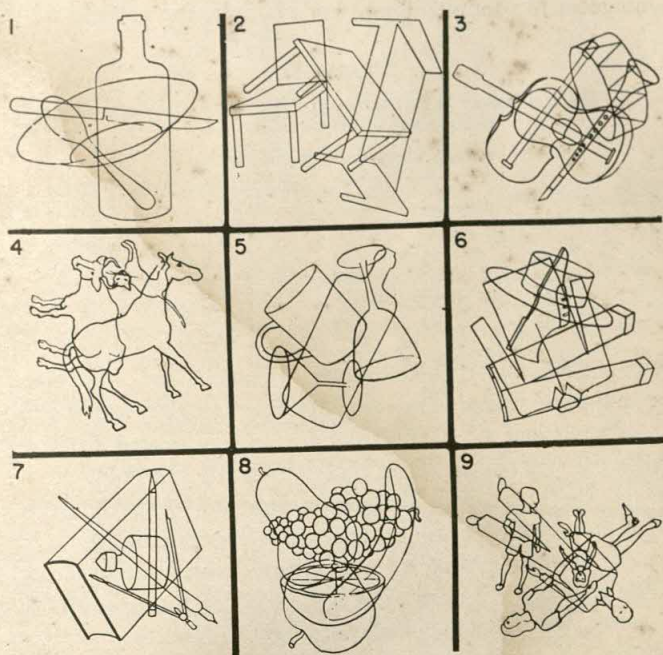
FIG. 1. OVERLAPPING, REALISTIC FIGURES
Each card was 4 in. on a side. The numbers did not appear on the cards used in testing.

Further, there were surprisingly few omissions at all age-levels, with the youngest group showing a median number of only 2 omissions among the 35 items. The number of omissions was a conservative measure of what *S* actually missed, since instances of misnaming were not taken into consideration. A high level of performance was still obtained, however, when the score used was the number of items correctly named or traced. Thirty of the 35 items were thus identified by the youngest group.

Discussion. The improvement in performance with age could be due to one or more of the following factors: greater familiarity with the names and types of objects; increased capacity to recognize line drawings; or in-

creased ability to unscramble overlapping figures. Analysis of the errors in response to individual items suggests that each of these possibilities was a factor at some time. First, the item most frequently omitted was the compass, and it is likely that the Ss were least familiar with this object and its name. Secondly, some of the misnamings indicated that the drawings were ambiguous to young children. For example, the orange was called a

TABLE I

DISTRIBUTION OF Ss ACCORDING TO NUMBER OF OMISSIONS FOR REALISTIC OVERLAPPING FIGURES

(Numbers represent the percentage of Ss within each group.)

Age (yr.)	N	Number of omissions					
		0	1	2	3	4	5
4-5	40	13	20	30	22	5	5
6-7	24	50	29	17	22	5	10
8-9	22	77	23	—	4	—	—
10-12	13	85	15	—	—	—	—

"basket" or "bowl" by 10 Ss, and some of the omissions of the orange may have been due to inability to recognize the line drawing. Thirdly, the number of omissions of a particular figure was related to its context, since the bottle was omitted eight times on Card 5 but only once on Card 1, although the bottle was identical in size and shape on both cards.

The results obtained thus far raise several questions. Was the good performance in unscrambling the overlapping figures due to the presentation of relatively familiar objects? Was the good performance related to the use of overlapping rather than embedded figures? Was the improvement in performance with increasing age due primarily to increasing capacity to recognize and to name line drawings? The experiment to be reported next was concerned with these problems. Comparisons were made between realistic and geometric overlapping figures, and between overlapping and embedded figures. To minimize the effect of ability to identify and name line drawings in producing improvement with age, a multiple-choice method of response was used with the geometric, overlapping figures.

EXPERIMENT II

Subjects. Thirty-four children from 4 to 8 yr. of age, who had not participated in the first experiment, were tested at the child care center referred to above. As can be seen in Table II, there were 8 Ss in each of 3 age-groups, and 10 Ss in a fourth group, with an equal number of boys and girls at each age-level.

Procedure. Three different series of pictures were shown to each child in the order described below.

TABLE II

DISTRIBUTION OF Ss ACCORDING TO NUMBER OF OMISSIONS FOR REALISTIC
AND GEOMETRIC OVERLAPPING FIGURES

(Numbers represent the percentage of Ss within each group.)

Age (yr.)	N	Type of figure	Number of omissions			
			0	1	2	3
4	8	realistic	12	25	25	38
		geometric	38	62	—	—
5	10	realistic	40	40	10	10
		geometric	80	20	—	—
6	8	realistic	50	50	—	—
		geometric	100	—	—	—
7–8	8	realistic	88	12	—	—
		geometric	100	—	—	—

(1) Series A consisted of overlapping, realistic figures. Cards 1, 6, 8, and 9 were selected from the nine cards previously described, and were presented to S in the orientation shown in Fig. 1. S was told to name, or show with his finger, all the things that he saw on the card; he could hold the card any way he wished.
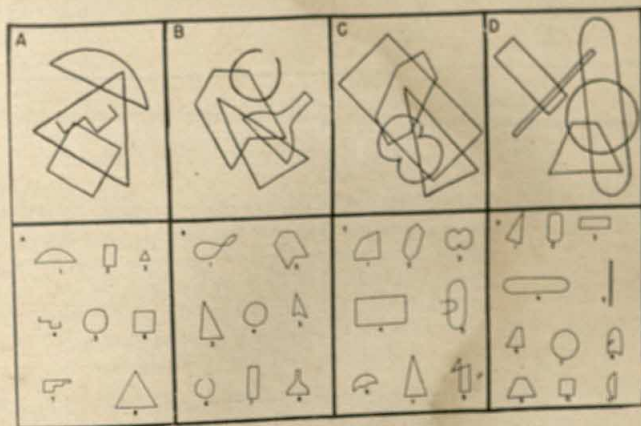
FIG. 2. OVERLAPPING, GEOMETRIC FIGURES

Each card of overlapping figures was 4 in. on a side; each sheet of multiple-choice items was 8 by 11 in. The letters did not appear on the cards used in testing.

(2) Series B consisted of overlapping, geometric figures, presented to S in the order and orientation shown in Fig. 2. Each card was shown with its corresponding sheet of multiple-choice items, among which could be found the figures on each card. S was told to find, on the large sheet, all of the pictures that he saw on the small card.

(3) Series C consisted of embedded, geometric figures from the Thurstone modification of the Gottschaldt figures, presented to S in the order and orientation shown

in Fig. 3.[9] S was shown the simple figure, and was told to trace the part that looked exactly like it in the complex figure. The first two figures were used for practice, with E showing S the correct response, and then allowing him to trace the figure himself. S was then given the rest of the cards, one at a time, and allowed as much time as he wanted to complete the task. The four test-figures used in Series C were identical in shape and size with one of the figures in each of the four cards of Series B.

Scoring. Two methods of scoring were used. Series A and B were compared for number of omissions. As previously described, an omission was complete absence
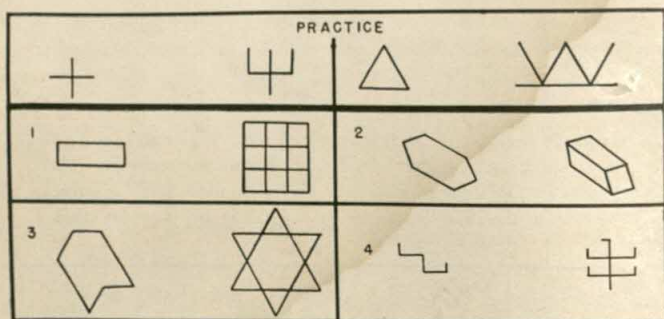
FIG. 3. EMBEDDED, GEOMETRIC FIGURES
Each card was approximately 8 by 2 in. The numbers did not appear on the cards used in testing.

of reference to a particular item; misnaming in Series A or mismatching in Series B did not constitute an omission. Series B and C were compared for number of errors. An error was any item incorrectly traced in Series C or any item selected from the multiple-choice in Series B that differed in shape or size from the corresponding item on the card of overlapping figures. In order to have the same number of items contributing to the score in Series B and C, only the four items were scored in Series B (one on each card), that were identical to the test-figures in Series C.

Results. For both series of overlapping figures, the number of omissions was small as can be seen in Table II. Each set of four cards contained a total of 17 items; the median numbers of omissions in the youngest group were only two and one of the realistic and geometric figures respectively. Although the difference was slight between the two series of figures, a significantly smaller number of omissions did occur for the geometric figures. Of the 20 Ss, who omitted any items on either set of figures, 3 Ss omitted an equal number of each set, 2 omitted more of the geometric figures, and 15 omitted more of the realistic figures. The observed propor-

[9] L. L. Thurstone, *A Factorial Study of Perception,* 1944, 72-75.

tion of Ss omitting more items of the realistic figures was significantly greater than the proportion expected on the basis of chance ($p = 0.01$).

While the number of omissions for geometric, overlapping figures was small at all ages, there was significantly better performance with increasing age. To test this observation with the small number of Ss at each age-level, the children of 4 and 5 yr. were combined into one group and compared to the group from 6 to 8 yr. with respect to the number of Ss omitting one item and zero items. Using chi-squared with Yates' correction, the difference between the age groups was significant at the 2% level.

To compare performance on the overlapping and embedded figures, error-scores were used instead of omissions. Since one form on each card in Series B was identical to one form on each card in Series C, the errors made on this form were compared in the two series. Table III shows the distribution of Ss in each group according to the number of errors made on the two types of figures. The median score of the 4-yr.-old Ss was less than one error on the overlapping figures, but almost four errors on the embedded figures. No S made more errors on the overlapping figures than

TABLE III

DISTRIBUTION OF Ss ACCORDING TO NUMBER OF ERRORS FOR OVER-
LAPPING AND EMBEDDED FIGURES

(Numbers refer to the percentage of Ss within each group.)

Age (yr.)	N	Type of figure	Number of errors				
			0	1	2	3	4
4	8	overlapping	38	38	25	—	—
		embedded	—	—	—	25	75
5	10	overlapping	70	30	—	—	—
		embedded	—	—	10	40	50
6	8	overlapping	88	12	—	—	—
		embedded	12	25	38	25	—
7–8	8	overlapping	100	—	—	—	—
		embedded	25	38	25	12	—

on the embedded figures; 3 Ss made no errors on either series. There was significant decrease in errors on embedded figures with increasing age. A fourfold table comparing age with errors shows a relation significant beyond the 1/10% level.

Preliminary work with the series of overlapping, geometric figures of Series B had indicated that most Ss responded to the large forms, *i.e.* those constituting the smallest number of forms necessary to describe accurately the picture. It is obvious, however, that the figures could be broken up

into smaller units. A few of the possible small-unit figures were included in the multiple-choice sheets. The younger Ss tended to pick these small-unit figures in addition to, or instead of, the larger forms. In the group of 18 Ss from 4 to 5 yr. old, 12 chose one or more of the small figures. In the group of 16 Ss from 6 to 8 yr. old, 3 made such choices. Using chi-squared, the difference between the age-groups was significant beyond the 1% level.

Discussion. There are two major findings. First, as in Experiment I, even the youngest Ss showed a high level of performance for the overlapping figures, with a small, but consistent, improvement appearing with increasing age. Secondly, there was a striking difference in the level of difficulty of embedded and overlapping figures.

The good performance with the geometric, overlapping figures is surprising in view of Osterrieth's report that a complex geometric figure is seen by young children in simplified form, with attention focused on details of the figure.[10] This conclusion, however, was derived from work with an embedded type of figure, which is a difficult task for young children. Thus it cannot be concluded that young children generally exhibit great difficulty with respect to pictures in which figure and ground are not clearly delineated. Rather, young children exhibit great difficulty in analyzing embedded figures, and relatively little difficulty in analyzing overlapping figures, regardless of whether they are realistic or geometric.

The question was raised earlier whether the improvement with age found in Experiment I was due to an increase in the ability to identify and name objects rather than to a change in the capacity to unscramble overlapping figures. With the geometric figures of Series B, S was merely required to match the forms, and hence the response was relatively independent of verbal skills or ability to recognize line drawings. Nevertheless, there was a significant improvement with age for these figures. It must be concluded that there is a change with age in the ability to *distinguish* figures that overlap each other.

To seek the factors underlying improvement in performance, one must first raise the question, "How are the figures differentiated?" From watching the Ss, it was clear that the discrimination occurred over a period of time, possibly more quickly in the older Ss. Most Ss moved their eyes over the figure, and it appeared to E as though S followed a group of adjacent lines until he found a configuration that assumed the properties of a figure.

---

[10] Osterrieth, *op. cit.,* 225-233, 254, 319-320.

As an hypothesis, then, it is suggested that the increase with age in ability to unscramble overlapping figures reflects an increase in perceptual span.[11] That is, an increase occurs in the number of lines that may be perceived simultaneously, and the number of lines that may be remembered after the eyes have shifted to another spot. Preliminary investigation with the use of a peephole for scanning the picture indicates that this method may be helpful in analyzing such a sequence of behavioral events in discriminating forms.

The good performance on both series of overlapping figures has been stressed because it was anticipated that the children would have difficulty, at least with the geometric figures. Actually, there was a slight, but significant, difference in favor of the geometric figures. The use of a multiple-choice method with the geometric figures could have increased the probability both of perceiving the figures and of responding to them. The perception of the geometric figures may have also been facilitated by prior experience in differentiating figures, since the geometric series was always presented after the realistic pictures. It is likely that variations in procedure were responsible for the slight difference in performance for the two series of overlapping figures.

The most striking result of Experiment II was the poor performance for the embedded figures, in contrast to the overlapping figures. This inferior performance cannot be attributed to variations in procedure since the variations were such as to favor the embedded figures. The embedded task was always presented last that S could profit from experience with geometric forms. For the embedded figures, unlike the others, two practice cards were used, hence S actually had experience with the problem. Furthermore, the mode of presentation of the task had even more advantages than those usually offered by a multiple-choice method, since S was required to find only one simple figure in the complex figure. Comparison between simple and complex figures was facilitated since they were both on the same card and in the same orientation. In the overlapping figures, comparison was not facilitated by proximity nor by similarity in orientation. It must be concluded that the large difference in performance for the overlapping and embedded figures was due to the greater difficulty in distinguishing the correct figure in the series of complex forms.

Variation in the difficulty of tasks involving the differentiation of forms from each other has been noted earlier. Vernon, in summarizing this work,

---

[11] This idea is similar to Piaget's concept of the increase in "the field of attention" as the child grows older. See Jean Piaget, *Judgement and Reasoning in the Child*, 1928, 218-232.

points out that two aspects of the task appear to be related to performance.[12] First, when the figure to be found has a "good" or "natural" structure of its own, it is easier to find than when it has an "unnatural" structure. Secondly, when the larger configuration forms a unified whole so that the figure to be found does not exist phenomenally, the figure is more difficult to find than when it is merely 'masked ,' but not 'submerged.' (The distinction between submerging and masking would seem to correspond to the description used here of embedded and overlapping.) Actually, the goodness of the figure and the inclusiveness of the configuration are both reflections of the phenomenon of grouping in visual perception. The basic problem is why lines are organized into the particular groupings that they are, and why the ease of selecting particular groupings under certain conditions changes with age.

Perhaps, however, we may sidestep the problem of organization in visual perception, and try to define what it is that the experimenter can do in order to make a particular form easy or diffcult to find. That is, is there any way of adding forms to a given figure such that the figure is easy to find in one case, and difficult to find in another case. The results previously described suggest one possible answer to this question. In the cards of overlapping forms, several forms were so combined that the contours intersected with each other, but they never coincided. In the cards of embedded figures, the configuration of lines was presumably perceived as a single form, and thus the figure to be found must have shared contours with the larger unit.

As an hypothesis, then, it is suggested that a figure is easy to find when other figures (certain arbitrary, predetermined groupings of lines) are so added that they intersect with the original figure and hence share points but do not share contours. The figure is hard to find when other figures are added so as to use, in their construction, one or more contours of the original figure. For example, a square can be 'hidden' in two ways, even though each method uses the same number and kind of added figures; namely, four triangles. The triangles can be so added that they intersect with the square and each other, but never share contours. On the other hand, the triangles can be so added that they share contours with the square and with each other. It might be possible to elaborate a continuum (or continua) of amount to which a given figure shares contours with other added forms. The minimum would be sharing of a point (intersection of lines), and then increasing numbers of contours could be shared with one other figure or increasing numbers of contours could be shared with separate

figures. The extreme case of contour-sharing with one other figure occurs in the reversible figure.[13] The extreme of contour-sharing with many forms does not lead to any predictable type of figure.

An attempt to study degrees of contour-sharing in relation to the difficulty of the task may quickly lead to a dead end, since sharing of contours can be produced in a variety of ways, and it may be that the different kinds of juxtaposition of forms would override any effects predictable on the basis of contour-sharing. We believe, however, that all ways involving contour sharing will hide a figure more frequently than ways involving contour intersection of the same figures. A prediction stated so broadly and without regard for principles of organization in visual perception must of course be wrong, but it can be investigated and qualified by empirical findings

If there is any validity to the idea that contour-sharing forms present special difficulty to young children, then it could be said that young children have difficulty in perceiving a given boundary as simultaneously belonging to more than one form. This line of thought is consistent with the suggestion made earlier that young children have a narrow perceptual span, *i.e.* a relatively restricted number of lines may be seen simultaneously, and remembered after the eyes have shifted to another spot. The capacity to see and remember a certain number of lines at any one moment would presumably have to reach a certain level of development before the child would be able to perceive a given line as belonging to more than one form.

Difficulty in perceiving a line as simultaneously belonging to more than one figure may be related to the observation that children sometimes respond to aspects of a picture that would not appear segregated to an adult. For example, with the geometric, overlapping figures, the 4- and 5-yr.-old children listed small forms in addition to, or instead of, the larger forms. One could simply say that children respond to small units, but this would not account for the instances where the children listed the large forms as well. It can be noted that the small forms did not completely overlap with the large ones, but shared contours with them to some extent.

At first thought, it seems contradictory that the younger children would list the small, 'embedded' figures more frequently than the older children, since it has been stressed that the younger children showed greater difficulty in discovering embedded figures. The contradiction is, however, more apparent than real. It has been suggested that the younger child's poor

---

[13] It is of interest in this connection that children under 8 yr. of age appear to have difficulty in seeing reversals on a Necker cube (R. K. Meister, Development of reversals in the Necker cube, *Research Relating to Children,* U. S. Department of Health, Education, and Welfare, Bull. II, Suppl. 2, 1952, 21). Direct evidence of a relation between performance on reversible figures and embedded figures in adults is reported by Newbigging, who found a positive correlation between correct scores on a series of modified Gottschaldt figures and number of reversals on several reversible figures (P. L. Newbigging, The relation between reversible perspective and embedded figures, *Canad. J. Psychol.,* 8, 1954, 204-208).

performance on embedded figures reflects difficulty in seeing a given set of lines as simultaneously belonging to more than one figure, and thus, the small figures were not seen as part of the large figures. It is likely that the older children saw the small figures just as frequently as the younger children, but the small figures were not reported because they were presumably seen as parts of the larger ones. It is predicted that the older children would actually be able to find more of the smaller embedded figures, if asked to do so, than the younger children. In fact, Osterrieth has reported that young children find fewer forms than older children when presented with a complex figure in which some subforms could be considered to be embedded in others.[14]

There is some evidence to suggest that the type of figure found to be more difficult for the children is also relatively more difficult for brain-injured Ss. Werner and Strauss found that brain-injured children correctly recognized figures, after an inspection of several seconds, which were partly covered by other lines but had great difficulty on the marbleboard task which involved the superimposition of a pattern of marbles on another pattern of holes, i.e. a task in which the patterns shared contours.[15] Of the brain-injured adults tested in this laboratory by Teuber, Battersby, and Bender on the Thurstone modification of Gottschaldt's figures, some men were also tested on the overlapping, realistic figures (Fig. 1).[16] In every case, performance was much better on the overlapping figures.

It is clear, then, that young, normal children, and perhaps brain-injured patients as well, do not exhibit great difficulty with respect to all pictures in which figures are not clearly set apart from each other. Poor performance cannot be attributed simply to a unitary and generalized deficit in ability to articulate figure and ground. It is suggested that a given group of forms would be differentiated relatively easily when the forms are so combined that their contours intersect, but never coincide, whereas the same forms would be differentiated with difficulty when they are combined so that their contours coincide.

SUMMARY

This investigation was concerned with the child's perception of figures that are not completely separated from each other. In the first experiment, a series of overlapping, realistic figures was shown to 99 Ss from 4 to 13 yr. of age. Very nearly all the intended figures were reported, and the

---

[14] Osterrieth, op. cit., 254, 296-298.
[15] Werner and Strauss, op. cit., 239 and 241-243.
[16] Eight men traced correctly 10 items or less (of the 61 items) on the modified Gottschaldt figures, whereas the same men responded correctly to 32 items or more (of the 35 items) on the overlapping, realistic figures.

number of figures omitted was small even in the youngest group. Omissions decreased with age.

In the second experiment, comparisons were made between realistic and geometric figures, and between overlapping and embedded figures, in a group of 34 children from 4 to 8 yr. of age. For the overlapping figures, both realistic and geometric, the number of omissions was again small even in the youngest group. Significantly more errors were made on the embedded figures than on the overlapping figures. Performance improved with age for all three series of figures.

It was suggested that when a figure was so 'hidden' by other forms that the boundaries of the added forms coincided with those of the original figure, the figure would be harder to find than when the boundaries of the added forms intersected with those of the original figure. If there is any validity to such an analysis of the child's difficulty with the embedded figures, then it could be said that the improvement with age reflects an increase in the capacity to perceive a boundary as belonging to more than one figure.

THE MODIFIED FECHNER-WEBER LAW AS A COMPLEX LAW OF DOSE-ACTION

By Albert Bachem, University of Illinois College of Medicine

Authors of textbooks of physiology and psychology still consider Fechner's law as a peculiarity of some of the special senses. Fulton, in particular, believes that the "endlessly criticized" law expresses a "fundamental feature of sense-organ behavior."[1] It is true that the law has been repeatedly criticized, interpreted, corrected, and explained. Many mistakes, however, have been made in these undertakings. The present contribution attempts to correct some of the mistakes and also to modify and explain the law on a broad physiological basis.

(1) Fechner's law. Fechner's law, in its simplest form, states that "the sensation increases with the logarithm of the stimulus."[2] The logarithmic formulation of the law is: $s = k \log i$. This law and its formulation have been questioned from numerous points of view, chief among them being: (1) the appropriateness of integrating an increment that does not converge to zero; (2) the attribution of equal numerical values to the consecutive thresholds, particularly of the absolute and differential limens; and (3) the assumption that sensations may be measured by arranging them in a quantitative order.[3]

(2) Criticisms not justified. These criticisms are, however, not justified because: (1) differential calculus has led to many useful results when the increments do not converge to zero, *e.g.* the quantum theory in physics; (2) there is no difference in principle between the addition of sensory units and the addition of energy or weight units; and (3) the belief that sensations and other psychological items cannot be measured like physical objects can only grow on a dualistic ground. Physics deals with nothing else but sensations, concepts, and universals. Measurement in itself is a psychological process. Moreover the quantitative psychological work, introduced by Fechner and Wundt and continued by modern experimental psychologists, has led to consistent results. The establishment of accurate pitch and intensive scales by multiple fractionation, the construction of loudness and isophonic contours, the measurement of psychological time through empty and filled intervals represent only a few examples of psychological accomplishment, not different in principle from physical measurements and scales.

(3) Fechner's interpretation. Fechner proposed his semilogarithmic law as the 'fundamental formula' of psychophysics, representing the quantitative relationship between stimulus and sensation. He emphasized that the logarithmic transformation could not conceivably occur between the physical and physiological acts, nor anywhere within the chain of physiological events between receptor and sensorium. He placed the transformation at the borderline between the physiological brain function and the psychological activity, and he postulated it as of mathematical accuracy. Deviations of the logarithmic formula towards both ends of its range were attributed to physiological factors, such as the intrinsic light of the retina and the dazzling effect of high intensities of illumination. This broadly designed attempt of Fechner must, as Von Weizäcker points out, be considered as a total failure.

(a) The famous measurements of König and Brodhun showed that the $\Delta i/i$-values were only approximately constant over a middle range of light intensity, but increased progressively toward smallest and largest intensities.[4] Knudsen, Riesz, and others made the corresponding observations for sound.[5] Biedermann and Löwit demonstrated the same relationship for weights,[6] Stratton, Gatti and Dodge, Kiesow, and others for touch.[7] According to all this evidence, Weber's law holds only approximately over a limited range.

(b) As Fig. 1 shows, the integration of Weber's function (Curve a) leads in Fechner's sense to the sensory Curve b with a threshold and a sigmoidal shape. The positively curved (concave upward) part corresponds to a declining $\Delta i/i$, while the negatively curved (concave downward) part corresponds to increasing quotients. The point of inflexion coincides with the minimal value of the quotient. The broader the minimum appears, the more a straight line is approximated on both sides of the inflexion. When Weber's function is plotted semi-logarithmically and then integrated, the straight-line region around the point of inflexion represents the original Fechner-law. When Weber's function is plotted on a linear scale the convex part of the resultant integration curve represents the approximate logarithmic relationship, as can be demonstrated by semi-logarithmic replotting.

(c) Most physiological dose-action curves contain sigmoidal and logarithmic elements, associated with fading-out or saturation-indications. A glance through A. J.

[4] A. König and E. Brodhun, Experimentelle Untersuchungen über die psychophysische Fundamentalformel in Bezug auf den Gesichtssinn, *Sitzsber. preuss. Akad. Wiss.*, 1888, 917-931; 1889, 641-644.

[5] V. O. Knudsen, The sensibility of the ear to small differences in intensity and frequency, *Phys. Rev.*, 21, 1923, 91-94; R. R. Riesz, Differential sensitivity of the ear for pure tones, *ibid.*, 31, 1928, 867-875.

[6] Biedermann and Löwit quoted by E. Hering, Zur Lehre von der Beziehung zwischen Leib und Seele, *Sitzgsber. Akad. Wiss. Wien., Math. Natur. Kl.*, 72 (3), 1875, 342-345.

[7] G. M. Stratton, Ueber die Wahrnehmung von Druckänderungen bei verschiedenen Geschwindigkeiten, *Phil, Stud.*, 12, 1896, 535-538; Alessandro Gatti and Raymond Dodge, Ueber die Unterschiedsempfindlichkeit bei Reizung eines einzelnen, isolierten Tastorgans, *Arch. f. d. ges. Psychol.*, 69, 1929, 405-426; Friedrich Kiesow, Zur Frage nach der Gültigkeit des Weberschen Gesetzes im Gebiete der Tastempfindungen, *ibid.*, 47, 1924, 1-13; Selig Hecht, Vision: I. The nature of the photoreceptor process, *A Handbook of General Experimental Psychology*, Carl Murchison, ed., 1934, 704-828; The visual discrimination of intensity and the Weber-Fechner law, *J. gen. Physiol.*, 7, 1924, 252.

Clark's *Mode of Action of Drugs on Cells*, under consideration of linear and semi-logarithmic plotting, will bear this out, as Table I demonstrates.

The outstanding exception to these dose-action curves is the all-or-none response
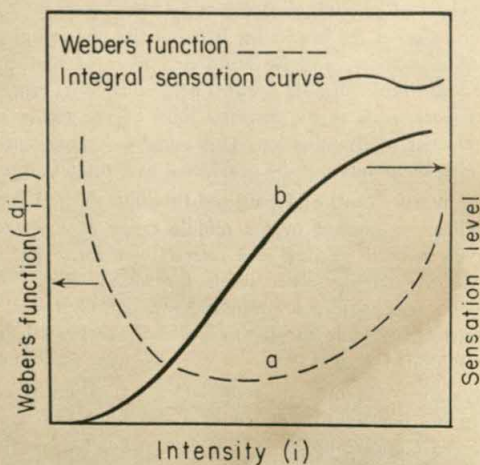
FIG. 1. INTEGRATION OF WEBER'S FUNCTION

TABLE I

FIGURES IN CLARK'S *Mode of Action of Drugs on Cells* (1933) ILLUSTRATING
SIGMOIDAL AND LOGARITHMIC ELEMENTS OF DOSE-ACTION CURVES

Function	Pages
True ogive*	15A, 98, 99, 126C$_1$, 128, 134A
Sigmoid	105, 111, 126C$_2$, 140C, 215
Sigmoid with fade-out	101, 137, 207
Approximately logarithmic (quasi-logarithmic)	102, 140A, 140B, 210 213, 218, 220, 221, 223
Quasi-logarithmic with positively and negatively curved endings	107, 113, 114, 144, 148, 152, 155 156, 157, 158, 255, 256, 258

* The term 'ogive' refers to a sigmoid which corresponds accurately to a Gaussian (symmetrical on a linear scale) probability distribution.

of muscle fibers, the cardiac syncytium, and some glandular elements which, however, do not appear as rigorous now as before).

The pharmacological and therapeutic literature contains a wealth of quantitative dose-action material (for epinephrine, histamine, ergotamine, acetylcholine, etc.) associated with hyperbolic, parabolic, exponential, and other formulas for practical use and for theoretical consideration. All these formulas discard minimal and maximal values of doses and actions. Semilogarithmic plotting usually reveals dose-action

curves composed of positively curved, logarithmic and saturation-parts which have more physiological significance, as will become evident later.

The fact that, with few exceptions, dose-action curves are composed of positively curved, logarithmic, and negatively curved elements, thus resembling the stimulus-sensation curves, rules out Fechner's psychophysical interpretation of the latter. It rather suggests a purely physiological mechanism, directly responsible for the psychological effect. With other words, the quantitative transformation between stimulus and sensation occurs within the physiological mechanism, not on the border between body and mind. Various factors have been made responsible for this physiological transformation by different authors.

(4) The excitatory process. Hecht has paid particular attention to the chemistry of the excitatory process, limiting himself to the realm of vision. He designed a photochemical process, composed of primary and secondary responses, which enabled him to explain, even quantitatively, the threshold phenomenon, visual acuity, adaptation, discrimination of intensity and other visual items. Hecht postulated a primary response of the monomolecular type, *i.e.* of a logarithmic intensity-relation. He pointed out at several occasions (in connection with his work on mya, with his theory of dark adaptation, with his discussion of the experimental results of Wolf and Crozier, and with the bleaching of visual purple) that the primary photochemical reaction follows a monomolecular course *i.e.* is proportional to the logarithm of the intensity of the light stimulus.[8] This assumption is in agreement with the fact that many chemical and biological processes show such a logarithmic relation. The most significant one concerning light is the logarithmic response of the photographic reaction. In addition to this assumption, Hecht postulated a probability distribution of sensitivity for the receptors. Increasing stimulation activates an increasing number of receptors, with the result of increased visual acuity and increased brightness sensation. It is evident that Hecht has made the mistake of overemphasizing one factor (the excitatory process) and discarding other factors.

(5) Receptor sensitivity. Houstoun has pointed out the significance of the quotient $i/\Delta i$. instead of Weber's quotient $\Delta i/i$.[9] He considered the reciprocal of Weber's quotient as the "power of the eye to discriminate brightness." Its value "becomes zero, not infinite, in the case of a stoneblind man." Houstoun further discovered that the plotting of König's $i/\Delta i$-values fits closely the Gaussian probability curve.[10] This fit holds, however, only for a logarithmic recording of light intensities. In his first publications on the subject, Houstoun tried to explain the logarithmic probability curve through a random variation of the thickness of the layer between the photoelectric production of electrons and the electronic activation of the cones. Since the absorption of electrons follows an exponential law ($i_0 = ie^{-\mu x}$), a linear relation must hold along a logarithmic scale ($\log i - \log i_0 = \mu x$). (In his later textbook he has omitted this explanation.[11]) He was so convinced of the probability

[8] Hecht, *op. cit.*, 1934, 716, 728, 738, 757.
[9] R. A. Houstoun, On Weber's law and visual acuity, *Phil. Mag.*, 8, 1929, 520-529.
[10] Houstoun, The visibility of radiation and dark adaptation, *ibid.*, 10, 1930, 416-432.
[11] Houstoun, *Vision and Colour Vision*, 1932, 15.

nature of the $i/\Delta i$ function that he looked for reasons for observed deviations from the normal probability distribution.

Houstoun saw several possibilities: (a) the Gaussian distribution is the exception rather than the rule. Most random distributions, occurring in nature are skewed due to one-sided variations of one factor; (b) the main system may be overlapped by a secondary; (c) spherical aberration could account for asymmetries; and (d) he also mentioned the possibility of exhaustion of the photosensitive substance at high intensities and the refractory period of the nerve fibers.

Houstoun fell into the same error as Fechner did insofar as he first believed in the absolute applicability of one principle, and then tried to explain its defaults. Actually the probability fit is best for white light. For blue light there is a pronounced deviation at low intensities; for red light the top of the curve is one-sided. Thus the symmetry of the whole curve appears as the result of the casual overlapping of several less symmetrical pure color curves.

The proper explanation of the broad logarithmic 'probability' curve is arrived at through observation of the resulting integral distribution-curve, which approaches a straight line over a considerable range, thus indicating an almost logarithmic function over an extended range.

(6) The all-or-none response. Hoagland makes the "all-or-none nature of nerve mechanisms" responsible "for the Weber-Fechner relation in the various sense departments."[12] He considers the "ogive curves, relating the amount of excitation to the logarithm of the stimulus" as "representing integral distribution-curves for thresholds of end-organs or afferent nerve-fibers." He emphasizes in particular that "the receptor units are regarded as coming into action in an all-or-none-fashion as their thresholds of sensitivity are reached on the curve of increasing intensity of light." The sensitivity of the receptors exhibits the normal probability distribution. The corresponding integral distribution curves represent true ogives (in Fechner's sense). Such are, in agreement with Hecht, not only the visual sensation-curves but also the visual acuity-curves plotted against light intensity. Hoagland quotes many dose-action responses of the ogival type, many of which are, however, plain random sigmoids. Thus he also makes the mistake of considering the sensory response as solely or mainly ogival in nature and attributing it to one factor, the all-or-none mechanism.

(7) Fiber-frequency. Besides the factors, emphasized and overemphasized by various authors, other factors have been neglected which must be considered as important for the sensory response. One is the individual fiber-frequency, as it increases with increasing intensity of the stimulus. This relation is, for low intensities, practically logarithmic as evidenced by semi-logarithmic plotting of the results of Adrian and Zotterman, Matthews, Hartline, Galambos, and others.[13]

[12] Hudson Hoagland, The Weber-Fechner law and the all-or-none theory, *J. gen. Psychol.,* 3, 1930, 351-373.

[13] E. D. Adrian and Yngve Zotterman, The impulses produced by sensory nerve endings, *J. Physiol.,* 61, 1926, 157; B. H. C. Matthews, The response of a single endorgan, *J. Physiol.,* 71, 1931, 73; H. K. Hartline, Intensity and duration in the excitation of single photoreceptor units, *J. cell. comp. Physiol.,* 5, 1934, 229; Robert Galambos and Hallowell Davis, The response of single auditory-nerve fibers to acoustic stimulation, *J. Neurophysiol.,* 6, 1943, 39-58.

All the curves, drawn from the results of Adrian and Zotterman on a single muscle spindle, begin as straight lines (with just noticeable positive or negative curving), but show a fading-out effect at higher intensities. Matthews' averaged curve from two experiments, also on the muscle spindle, is a perfectly straight line. Hartline's curve for initial frequencies of single photoreceptors is practically a straight line; the one for maximal frequencies shows the fading-out effect at higher
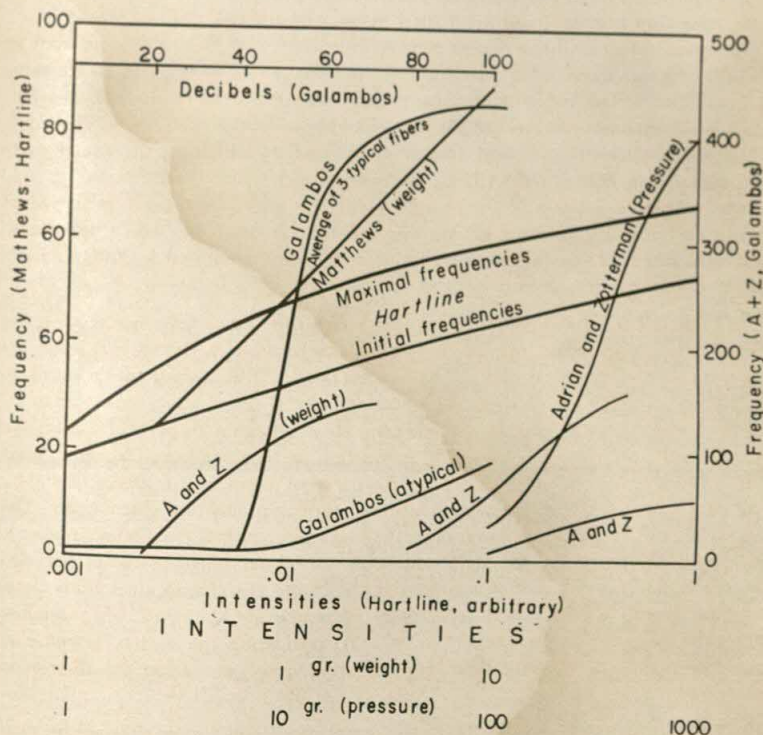
FIG. 2. SINGLE FIBER-FREQUENCY AS FUNCTION OF INTENSITY OF STIMULUS

intensities. Galambos' averaged curve from 3 typical auditory fibers is a practically straight line up to frequencies of about 300 impulses per sec. One 'atypical' fiber gives a perfectly straight line from 12 to 90 impulses per sec.

The logarithmic order of single fiber-frequency became evident through an interesting controversy concerning touch-discrimination. The work of Gatti and Dodge and of Kiesow, including the discussion of the experiments of Hansen, Stratton, and other independent observers, demonstrates that Weber's quotient is practically constant when single (or few) receptors are stimulated, but that the quotient shows a strong initial decline, when an extended or even expanding area is stimulated. In the first case, sensation increases logarithmically through the increasing number of

nerve impulses; in the second instance the sensation increase begins with a positively curved element through the increasing number of activated receptors.

These results indicate clearly a practically logarithmic relation between stimulus and single fiber-frequency, which becomes disturbed through the refractory period of the nerve fiber.

(8) Refractory period. In the relatively refractory period the nerve-ending requires a stronger than normal stimulus for the formation of the impulse and the frequency falls behind the logarithmic response. In addition, the rate of conduction decreases and the impulses arrive at the sensorium with a delay which reduces the frequency of its excitation.[14] The totally refractory period sets a final limit to the frequency-response of the nerve fiber. Thus the refractory period is responsible for the fading-out and saturation-effects at high intensities. As soon as all fibers discharge at maximal frequencies, the fading-out effect is complete and a maximal level of sensation is reached.

The fact that the increase of Weber's function toward higher intensities does not occur when the sense organ is well adapted through a reduced excitatory process and that in the corresponding integral sensation curve no fading-out takes place is highly significant.[15] It proves that the fading-out is mainly caused by an overload of the nerve fibers due to an excessive excitatory process.

(9) Other neurophysiological factors. It must be expected that the sensory response becomes further complicated through additional neural mechanisms. *(a)* Adaptation limits the response either through a restriction of the excitatory process or through a protective reflex mechanism. The most studied and best known excitatory process is the bleaching of visual purple. The fact that neither rhodopsin nor visual white, but an intermediate product (indicator yellow?) excites the rods, accounts for a stimulus-response which starts at a low threshold, reaches a maximum for dim light and drops out for strongest illumination. It seems that scotopic and photopic visions differ in principle concerning response to strongest stimulation. Also in comparison with other sensory mechanisms, scotopic vision appears as an exception to the rule. Unfortunately we know less about the normal excitatory processes than about the scotopic mechanism.

The outstanding adaptations by means of reflex mechanisms are pupillary constriction and stretch of the tympanum. Pupillary constriction protects against strong light, thus producing a fading-out effect. The maximal effect of normal response reduces the infalling light in a ratio of 16 to 1. König and Brodhun have eliminated this factor through a pin-point pupil, thus avoiding the fading-out effect. The reflex adjustment of the ear drum increases the sensitivity of the ear for high tones and decreases it for low tones of great intensity. In principle this mechanism is protective under conditions prevailing in nature. Generally speaking, gradual adaptation to increasing stimuli adds to the fading-out response.

---

[14] Francis Gotch, The delay of the electrical response of nerve to a second stimulus, *J. Physiol.*, 40, 1910, 250-274; H. S. Gasser and Joseph Erlanger, The nature of conduction of an impulse in the relatively refractory period, *Amer. J. Physiol.*, 73, 1925, 613-635.

[15] Hecht, *op. cit.*, 1934, 765.

(b) Central synapses account for absolute thresholds, since temporal and spatial summations are needed to overcome synaptic resistance. Spatial summation is affected by the conditions of sub-liminal fringe and occlusion (which are not limited to the reflex arc), the former predominating at weak activation, the latter at strong activity.

Attention is the psychic complement to a well-established thalamo-cortical conduction free from disturbing neural circuits. It promotes sensitivity and acuity at low intensities.

Masking of one sensory quality by an additional one at high intensities accounts for a decrease in discrimination of intensity, resulting in an end-rise of Weber's function and in a saturation effect of the Fechner-function. This was well observed in case of taste[16] and somesthesis,[17] and was noticed also in vision. Inhibition (and extinction) occurs through suppression-areas, particularly at high intensities. All these factors effect a steepening of the response curve near the threshold and a leveling-off at the upper end.

(10) *The statistical theory*. Crozier conducted extensive experimental and theoretical work on the differential threshold.[18] He criticizes Weber's law, Hecht's photochemical theory, and the use of formulations without "real analytical significance." He assumes, like Houstoun, that sensation results from "a frequency distribution of elements of neural effects in terms of log *i*." He clearly refers to neural activity or number of impulses in the part of the brain where experience takes place, "and not of the intensity-thresholds of excitable neural units." He gives, however, no convincing explanation for the logarithmic probability-distribution, nor does he explain the quasi-logarithmic character of the sensory curve, except that he calls for a maximal frequency of impulses in any particular number of neurons no matter how intense the stimulation.

The writer considers Crozier's postulation of a probability distribution of neural excitatability at the highest brain levels as an oversimplification. He also points to the fact that logarithmic transformation and saturation are accounted for at lower, peripheral levels. The differences between the sensation-curves of various modalities, as shown in the next paragraph, are not explained by the statistical theory either and require more individual considerations.

(11) *Comparison of various sensory responses*. A strict comparison of the response-curves of the various senses is difficult for several reasons. (a) Directly constructed response-curves exist only for vision and audition. They are obtained through fractionational experiments of brightness and loudness. (b) Indirect sensory response-curves can be integrated from Weber's function for all sensory modalities except

---

[16] F. Lemberger, Psychophysiche Untersuchungen über den Geschmack von Zucker und Saccharin (Saccharose und Krystallose), *Arch. f. d. ges. Physiol.*, 123, 1908, 293-311.

[17] A. H. Holway and C. C. Pratt, The Weber ratio for intensive discrimination, *Pyschol. Rev.*, 43, 1936, 322-340.

[18] W. J. Crozier, On the law for minimal discrimination of intensities: IV. Δ_i as a function of intensity, *Proc. Nat. Acad. Sci.*, 26, 1940, 382-389; S. S. Stevens and Hallowell Davis, *Hearing: Its Psychology and Physiology*, 1938, 89.

temperature and pain. (c) For the establishment of Weber's function different methods have been employed, *i.e.* threshold determinations and scatter responses for successive stimuli, flicker observations on vision, masking experiments on audition, point-like and extended stimuli for touch, etc. (d) Comparisons between different physical measures of adequate stimuli such as photons, dynes, and molar concentration are difficult and may be meaningless from the psychological point of view. Fechner's establishment of "the fundamental value of the stimulus" has overcome this difficulty. Holway and Pratt have applied this concept to comparative plotting of various Weber-functions.[19] The present writer has integrated the latter data into
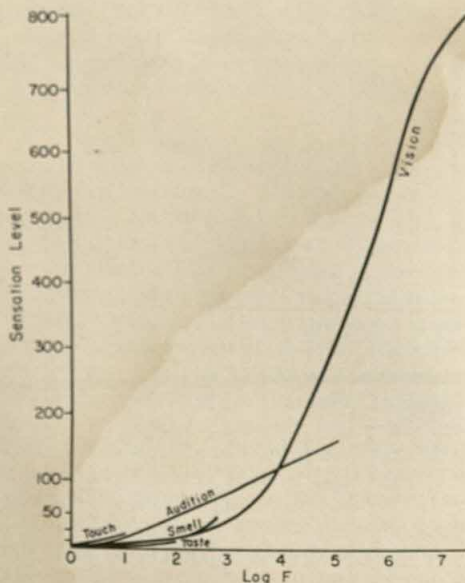
FIG. 3. COMPARISON OF FIVE SENSORY MODALITIES

integral sensory curves. He hopes that in spite of all the difficulties a relatively good comparison is guaranteed.

Fig. 3 compares five sensory modalities as to the stimulus-range over which they respond and to the sensory range which is undisturbed by additional qualities. The abscissa gives the logarithms of the stimuli in terms of Fechner's fundamental values (log F). The ordinates show the corresponding sensation-levels. The superiority of vision and the inferiority of taste and touch as to the range of discrimination and response become evident from this type of plotting.

(a) *Vision.* Fig. 4 presents the visual sensory curve established directly by averaging Houstoun's nine individual curves and those integrated from the Weber-function,

[19] Holway and Pratt, *op. cit.*, 355 ff.

averaged by Holway and Pratt from various observers.[20] The curves are definitely sigmoidal, with a central logarithmic approach over several logarithmic units. This fact indicates that receptor number, single fiber-frequency and fiber-refractoriness contribute noticeably to the brightness function.

In principle the electro-retinogram ('b'-wave and 'on'-response) and the opticus-
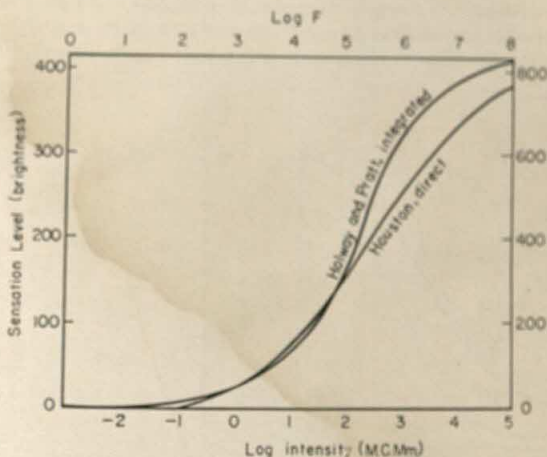
Fig. 4. Visual Sensory Curves

response are very similar to the sensory response of the eye. An exact quantitative comparison is impossible on account of different experimental conditions and incomparable stimulation-intensities. The similarity of the neurophysiological and psychological responses is particularly significant from the point of view, that retina and opticus are outposts of the brain (evagination of diencephalon) thus exhibiting central responses, such as summation, inhibition, and integrative scanning. The pupillary reflex, as studied by Hecht and by Houstoun also compares in general with the sensory response.[21] Such a close quantitative relationship between sensation and reflex supports strongly the interpretation of the corrected Fechner-law as a dose-action formula. (See Fig. 5.)

(b) *Audition.* The auditory, sensory curve (Fig. 6), derived from Weber's function, averaged from Holway and Pratt from many observers, is logarithmic over three to four logarithmic units. Only in the lowest range a positive curvature is noticeable.

The loudness curve, obtained through fractionation by several experiments, averaged by Stevens and redrawn on a semilogarithmic scale, shows a positive curvature practically over its whole range. This indicates that, for every comparison, the stronger tones are over-estimated somewhat in terms of bels or decibels through the

[20] Houstoun, *op. cit.,* 1932, 35; Holway and Pratt, *op. cit.,* 336.
[21] Hecht, *op. cit.,* 1924, 245; Houstoun, *op. cit.,* 1932, 10.

whole range. On a linear scale the stronger tones would be underestimated considerably—as indicated by Stevens double-logarithmic curve. (The relative position of the two curves along the abscissa is somewhat arbitrary.)

None of the two-type-curves exhibits a fade-out at the upper end. The middle-range discrepancy between the curves integrated from Weber's function and those obtained through fractionation is difficult to explain and may render one of the methods questionable. As a whole it appears that loudness is a combined function of receptor-number and fiber-frequency. This conclusion is supported by Galambo's observation on single fiber-responses over an extended frequency range.[22]
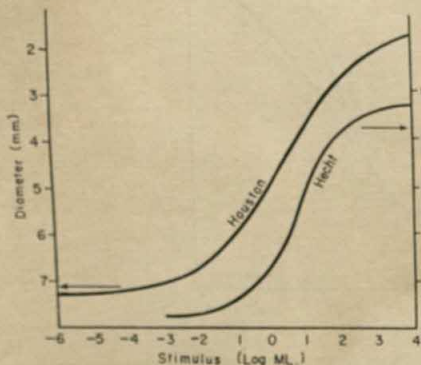
FIG. 5. CURVES OF PUPILLARY REFLEX TO LIGHT   FIG. 6. AUDITORY SENSORY CURVES
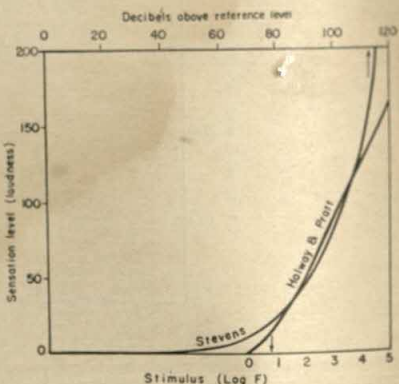
(c) Smell. The olfactory, sensory curve (Fig. 7) exhibits positive curvature throughout its complete range of about three logarithmic units. No logarithmic and saturation-responses are observable. This exceptional behavior of olfaction may be due to two peculiarities of the olfactory mechanism. (i) The lack of accessory, secondary receptor cells and of the associated excitatory process which is partly responsible for the logarithmic response. (ii) The existence of recurrent, amplifying circuits, in which the single fibers are not as much loaded or overloaded as in the predominantly parallel circuits of the other sensory mechanisms.

(d) Taste. The integrated curve for gustatory sensation (Fig. 8) extends over two logarithmic units and consists of a long positively curved part and short logarithmic and fade-out elements. The latter part coincides with the appearance of masking through touch (syrup-like sensation) and common chemical sensations (sharp, stinging taste).

(e) Touch. The integrated curve for touch i.e. cutaneous pressure (Fig. 9) approaches the logarithmic function more closely than any other sensory curve, particularly if a small area is stimulated and if the effects of deep pressure and pain are prevented. In this case the curve extends over one logarithmic unit only. The

---

[22] Galambos and Davis, op. cit., 47 f.

short, positively curved start indicates the increasing number of responding receptors; the fading-out signals the beginning of masking through alternate qualities.

(12) Comparison with dose-action curves. For a comparison of the sensory integrational and the dose-action curves, the author has established accurate dose-
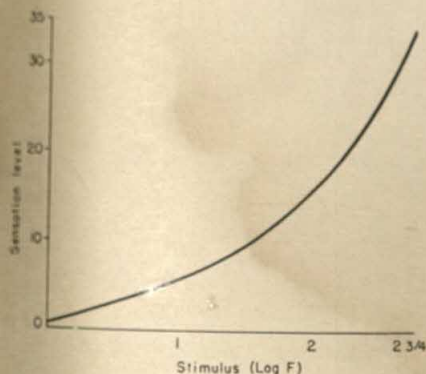
FIG. 7. INTEGRATED OLFACTORY SENSORY CURVE
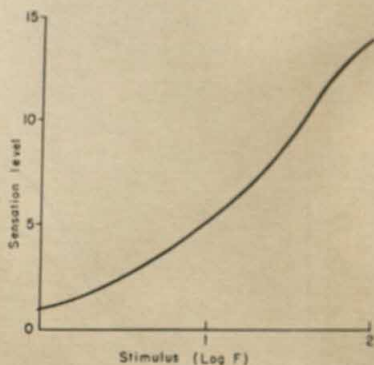
FIG. 8. INTEGRATED GUSTATORY SENSORY CURVE
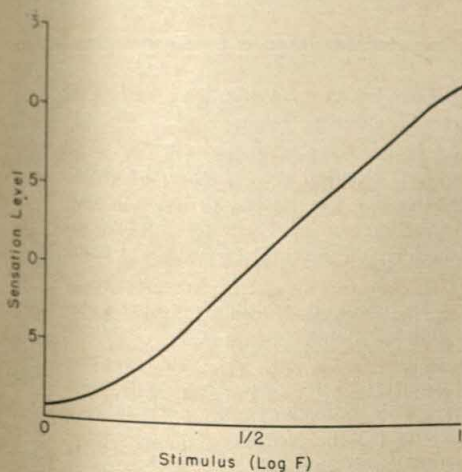
FIG. 9. INTEGRATED TACTILE SENSORY CURVE
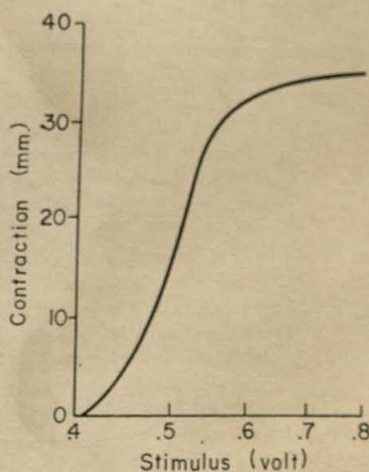
FIG. 10. MUSCULAR RESPONSE-CURVE

action curves for a few simple physiological effects, which were not available in the literature.

The electrically induced single muscle-twitch (Fig. 10) follows a sigmoidal dose-action curve, which extends over less than one half logarithmic unit between threshold and complete saturation. (Linear and semilogarithmic plotting makes little

difference over such a short stimulus-range, except that the start of the positively curved line is emphasized on the semi-logarithmic scale.) The logarithmic element is missing. The sigmoid appears as the integration of the somewhat skewed probability curve of the sensitivity of the various motor units, each of which reacting in an all-or-none fashion.

The same holds true for pancreatic and salivary secretions, evoked through injections of secretin and pilocarpine respectively (Fig. 11). Here the dose-action sigmoid is the result of the variable sensitivity of the glandular elements. The range
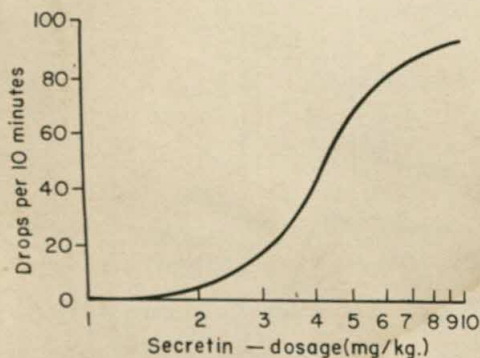
FIG. 11. PANCREATIC RESPONSE-CURVE   FIG. 12. SALIVARY RESPONSE-CURVE
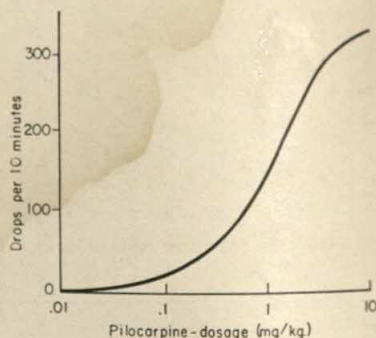
between threshold and saturation is more extended—1 to 3 logarithmic units—indicating a greater variation of the secretory elements. Sigmoids over 1 to 2 logarithmic units were also observed on the gastric response to histamine.[23]

When salivary secretion was evoked by a 5-sec. electrical stimulation of the corda tympani with increasing voltage, a definite saturation was reached at 2 v. A.C., indicating that from this voltage on all nerve-fibers responded maximally to the 60 v. cycle (Fig. 13). A quasi-logarithmic approach of the maximal response was here noticeable.

This interpretation is supported by an experiment, in which a constant 2 v. potential was applied for a variable time, with the result of an increased secretion for an increased product of voltage and time.

It is evident that the sensory response contains by far more of a logarithmic element than most other dose-action functions. This predominance of the logarithmic response is due to several physiological factors: (a) the excitatory response of the receptors, which is logarithmic, before saturation is approached; (b) the single fiber-frequency, which is logarithmic between the absolute threshold and the onset of relative refractoriness; (c) the overlapping of other, mostly neurophysiological factors, mentioned before.

Several dose-action curves approach the stimulus-sensation curves closely. The

[23] N. E. Hanson, M. I. Grossman, and A. C. Ivy, Doses of histamine producing minimal and maximal gastric secretory responses in dog and man, *Amer. J. Physiol.*, 101, 1932, 156.

pupillary light reflex was mentioned before in connection with the visuo-sensory response. Many responses to epinephrine are of the quasi-logarithmic type, *e.g.* the response of the nictitating membrane and of blood pressure. Rosenblueth[24] and
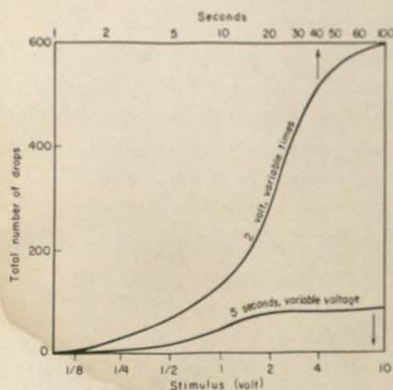
FIG. 13. SALIVARY RESPONSE-CURVES FOR VARIABLE POTENTIALS AND TIMES OF STIMULUS
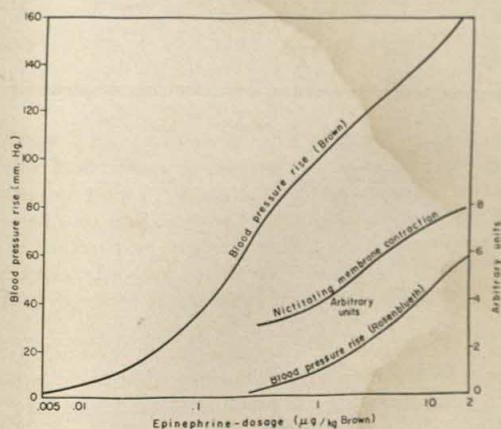
FIG. 14. DOSE-ACTION CURVES RESEMBLING STIMULUS-SENSATION CURVES

Brown,[25] discarding minimal dose responses, approximate the response curves by a hyperbola, but they consider the semilogarithmic curve as the second-best fit. A semilogarithmic plotting (Fig. 14) shows the great similarity of these dose-action curves with the sensation-curves. In these cases the physiological mechanism differs

[24] A. Rosenblueth, The mode of action of adrenin and the quantitation of adrenin by biological methods, *Amer. J. Physiol.,* 101, 1932, 156.
[25] R. V. Brown, A comparison of several dose-action curves for the pressor action of epinephrine, *J. pharmacol. exp. Ther.,* 98, 1950, 420.

somewhat from the sensory mechanism. Since receptors and excitatory response are missing, the smooth or cardiac muscle responds directly to the chemical stimulus, and a regulatory homeostatic, autonomic mechanism controls the final response. A quasi-logarithmic element results here from the overlapping of probability, fiber-frequency, and integrating brain-mechanisms.

Several auditory responses, though not of a dose-action type, are very closely logarithmic in nature. The relationship of frequency and tone-height, the so-called pitch-function, as determined by Stevens and Davis, is logarithmic within a range of two to three octaves, i.e. over almost one logarithmic unit.

The most peculiar, strictest logarithmic response is the frequency-chroma relationship observable in the genuine type of absolute pitch. This relationship is characterized by an octave periodicity, which can be represented mathematically by the logarithm to the base 2 of the frequency. The chroma-function is mathematically accurate between 30 and 4000~, but deviates somewhat below and decidedly above these limits, where complete chroma-fixation takes place, comparable to consecutive fading-out and saturation-effects.[26]

The frequency-responses of the ear depend upon four physical factors, i.e. length, tension, weight of the basilar fibers, and height of the lymph (endo- and peri-) column between the base and the response-area of the basilar membrane. All these factors collaborate harmoniously toward the establishment of the particular pitch-functions.

In contrast to the auditory frequency-response the visual response, resulting in the color spectrum, appears highly arbitrary, insofar, as an extended red region is followed by a series of narrow orange, yellow, green, blue and violet areas, and as a reversal toward blue takes place within the so-called near ultraviolet region.[27]

CONCLUSION

The sensory response is a complicated dose-action function, identical in principle with complex physiological responses. It exhibits absolute threshold and a sigmoidal form, characterized by postively curved, linear, and negatively curved elements on a semilogarithmic scale. The logarithmic element predominates definitely. Insofar, there is some truth in Fechner's and Fulton's statements, but Fechner's original psychophysical interpretation of his fundamental law on a mathematical, parallelistic basis is wrong. The recent attempts of explaining the sensory response through a single physiological factor (excitatory response, receptor sensitivity, all-or-none response of nerve fibers, probability function of cortical excitability) are erroneous, too. The sensory response is the result of the combined effects of many peripheral and central neurophysiological factors.

The absolute threshold is more of central than peripheral origin. It

[26] Albert Bachem, Chroma fixation at the ends of the musical frequency-scale, *J. Acoust. Soc. Amer.*, 20, 1948, 704-705; Tone height and tone chroma as two different pitch qualities, *Acta Psychol.* 7, 1950, 80-88.
[27] Bachem, The color of ultraviolet light, this JOURNAL, 66, 1953, 251-276.

represents the point at which synaptic resistance is overcome through temporal and spatial summation and an arousal of the sensorium is effected. The logarithmic response is mainly due to the excitatory process of the receptors and to the nerve fiber activity. The pre-logarithmic approach (positive curvature) coincides with the upward swing of the probability curve of the receptor-sensitivity and with the subliminal fringe of summation. The post-logarithmic fade-out coincides with the descending limb of the probability curve of the receptor-sensitivity, the relatively refractory period of the nerve fibers, masking, reflex adaptation, and central occlusion and inhibition. The saturation-effect is mainly due to the absolutely refractory period of the nerve fibers. All these effects overlap in quantitatively different ways for the different sensory modalities and qualities.

It appears possible that the sensation-level is exactly proportional to the neural activity, directly responsible for sensation, *i.e.* the multiple reflection of impulses (scanning) between the projection-areas, the centers of association and the centrencephalon. Such, a simple relationship between physiological and psychological elements may comfort the mind of behaviorists, isomorphists, psychophysical monists and other -ists. The writer chooses to abstain from theorizing on these issues at this time.

SUMMARY

The modified Fechnerian law represents a particular, complex dose-action law, characterized by absolute threshold and quasi-logarithmic response, and resulting from the interplay of many neurophysiological elements, *i.e.* excitatory process, distribution of receptor-sensitivity, responses of single nerve-fiber, synaptic summation, reflex adaptation, masking, central inhibition, and suppression.

STABILITY OF CHOICES AMONG UNCERTAIN ALTERNATIVES

By William H. McGlothlin, The Rand Corporation

Recently there has been an increasing interest in the theory of games and decision making, with the development of various models and strategies for determining choices. When decisions are made among alternative offers whose outcomes are unknown at the time of the choice, as in a gambling game, three main variables are involved. These are: (1) the amount wagered or risked, x; (2) the size or value of the prize, y; and (3) the objective probability of a successful outcome, P. If it is hypothesized that the individual's best strategy is to maximize the expected value, E, of his choices, it becomes a simple matter to predict uncertain decisions. Expectation, E, is defined as the summation of the products of all possible outcomes and the probability attached to each. Losses are treated as negative outcomes.

Human behavior often does not follow the above strategy, however, as in the case of buying insurance, or in accepting a gamble in which the expectation is negative. It may be assumed that the individual reacts to the psychological counterparts of x, y, and P; these are: utility of bet, $U(x)$; utility of prize, $U(y)$; and subjective probability, P'. Edwards and others have hypothesized that choices between uncertain alternatives can be predicted on the basis of maximization of 'subjectively expected utility,' SEU.[1]

$$SEU = \Sigma P'_i \, U_i$$

where U_i represents the utility of the ith possible outcome of the bet and P'_i represents the subjective probability of the outcome.

There have been some experimental attempts to measure utilities and subjective probabilities as functions of the corresponding objective scales. The validity of these functions is often in question, for it is usually nescessary to assume objective probabilities, or linear functions thereof, in order to find a subjective utility function, and to make similar assumptions about utility when deriving subjective probability. For instance, suppose that 75% of the time subjects (Ss) preferred

[1] Ward Edwards, Probability-preferences in gambling, this JOURNAL, 66, 1953, 349-364.

a probability of 0.2 of winning $4.00 to a probability of 0.4 of winning $2.00. From these results we could conclude one of the following: (1) psychologically, an objective probability of 0.4 is less than twice the objective probability of 0.2; (2) the utility of $4.00 is more than twice the utility of $2.00; (3) neither subjective probability nor the utility of money, as valued by S, can be expressed on the corresponding objective scales.

Instead of presenting S with alternatives yielding equal expected values as in the above example, we may design the experiment such that choices are made between events yielding unequal expected values. By using some sort of competitive bidding, or by means of an information processing center such as a pari-mutuel betting machine, it is possible to have S himself establish the expected values of various combinations of probabilities and prizes. Negative expected values for a particular bet, probability, and prize would indicate at least one of the following: (1) $P' > P$; (2) $U(\$x) < x$ utiles; $U(\$y) > y$ utiles; where utiles refers to the unit of measurement for utility. Positive expected values for a given set of values would indicate the reverse relationship; i.e. (1) $P' < P$; (2) $U(\$x) > x$ utiles; (3) $U(\$y) < y$ utiles.[2]

PROBLEM

The present study assumes there is sufficient comparability among individuals to make an investigation of group risk-taking behavior meaningful. It is a statistical study of 9605 thoroughbred horse-races, mostly from California tracks, during the years 1947-1953. The primary purpose is to examine the stability of risk-taking behavior over a series of events. One approach to this question is to determine the expected values of constant-size wagers for a range of probabilities of success P. This yields an E-vs.-P pattern and can be repeated for a series of risk-taking events, i.e. races. The stability of this pattern throughout the racing day allows some inferences to be made about the stability of subjective probability and utility for wager and prize over a series of events. It is also possible to obtain information about stability of risk-taking behavior that is independent of the expected-value. Variability of size of average wager over a series of events can be determined as well as preferences among wagers having equal expectations (Es) but different probabilities of success. Some limited information is available concerning differential preferences between winning and losing bettors.

PROCEDURE

Pari-mutuel wagering. Betting on horse-races is quite different from most other forms of gambling. The betting population establishes the odds, or the amount of money each horse in the race will return if successful. At a race track there is a

[2] For a more complete discussion of the problems involved in this type of model construction, see R. M. Thrall, C. H. Coombs, and R. L. Davis (eds.), *Decision Processes*, 1954, 255-285.

totalizator board on which the current odds-to-win on each horse appear. The odds are recalculated and flashed on the board at about 45-sec. intervals, keeping the public informed as to the amount of backing each horse has received up to that time. It is not until the final bet is made that the exact odds on each horse are established.

Betting may be for place or show in addition to win. If the horse finishes first or second the place ticket is redeemable; for the show ticket, the horse may finish first, second, or third. The amount returned on these tickets is independent of the position the horse occupies at the finish.

Win. The winning odds posted by the track are given by the formula: $a_i = [(1 - t) \cdot \Sigma A - A_i]/A_i$, where $a_i =$ odds that the ith horse will finish first; $t =$ proportion track takes;[3] $\Sigma A =$ amount bet in the win pool on all horses for the race being considered; and $A_i =$ amount bet on the ith horse to win.

The odds found in this manner are rounded downward to the nearest multiple of 5¢ (10¢ in some states). The odd cents so deducted are called breakage. The winning ticket pays an amount that includes these odds plus the original bet.

Place. Place odds at the track are determined by the following formula: $b_1 = [(1 - t) \cdot \Sigma B - (B_1 + B_2)]/2B_1$, where $b_1 =$ place odds for horse finishing in the first position; $\Sigma B =$ amount bet in the place pool on all horses for the race being considered; and B_1, $B_2 =$ amount bet that the horses finishing in the first and second positions will place. The place odds for the horse finishing in the second position, b_2 are found by replacing B_1 with B_2 in the denominator of the above formula.

Show. Show odds are determined as follows: $c_1 = [(1 - t) \cdot \Sigma C - (C_1 + C_2 + C_3)]/3C_1$, where $c_1 =$ show odds for horse finishing in the first position; $\Sigma C =$ amount bet in the show pool on all horses for the race being considered; and C_1, C_2, $C_3 =$ amount bet on the horses finishing in the first, second and third positions to show. The show odds for the horses finishing in the second and third positions, c_2 and c_3, are found by replacing C_1 with C_2 and C_3, respectively, in the denominator of the above formula.

It is apparent from the last two formulas that the place and show odds are dependent not only on the amount of money bet on the individual horse in these categories, but also on the amount bet on the other horses that appear in the numerator. While winning odds are calculated and reported on all horses in a race whether they win or not, it is practical to report the place odds only for the first two horses, and the show odds for the first three.

Range of odds. The range of odds established on the horses in any given race depends primarily on how closely the horses are matched in ability. In a typical race of 9 or 10 horses the odds-to-win range from around 2–1 on the public favorite to around 50–1 on the horse receiving the least public backing. The place odds typically range from about 1–1 to 20–1, while the show odds range from about 0.5–1 to 6–1. Thus, in a 9 horse race there are 27 possible bets with a typical range of 0.5–1 to 50–1. In the case of win-betting, good approximations of these odds are available to the bettor at the time he makes his choice. As explained above, accurate estimates of place and show odds are not available at the time of the decision making, although it is virtually certain that the place and show odds will be con-

---

[3] Track take varies from 10 to 15% in the 23 states permitting pari-mutuel wagering on thoroughbred racing. In California the figure is 13%.

siderably lower than the winning odds appearing on the totalizator board for a given horse.

Data.[4] The data used in this study were obtained from the *Daily Racing Form Chart Book* and are described in Table I.[5] The main sample consists of 1156 days or 9248 races. In view of the fact that some of the most interesting results were found in the data for the eighth race, an additional sample of 357 eighth races was analyzed to increase the reliability of the results. Whenever the eighth-race data were combined with data for other races, they were given a weight of 1156/1513.

Some tracks schedule an additional race on Saturdays, giving a total of nine races.

TABLE I

SOURCES OF DATA

Track	Years	Number of racing days
Hollywood Park, California	1947–1953	345
Santa Anita, California	1947–1953	348
Tanforan, California	1947, 1949–1951, 1953	212
Golden Gate Park, California	1947, 1949–1952	210
Bay Meadows, California	1951	41
Bay Meadows, California	1947–1950	168*
Jamaica, New York	1950	60*
Aqueduct, New York	1950	38*
Belmont Park, New York	1950	52*
Empire City, New York	1950	5*
Saratoga, New York	1950	34*
	Total	1513

* Only eighth races included.

To combine these races with the remainder of the data, the fifth race was omitted and the sixth race was used in place of the fifth-race data, and so on.

The study is of a statistical nature, and as such, lacks many of the controls found in the experimental laboratory setting. The population of bettors is not stable throughout the racing day due to late arrivals, early departures, and the fact that many bettors do not wager on every race. The amount bet by different individuals varies in an uncontrolled manner, such that persons wagering large amounts determine the size of the odds to a greater extent than do smaller bettors. Also, there is no direct way of studying differential behavior among those persons receiving reinforcement in the form of successful bets and those losing. Finally, the results found are strictly applicable only to the population from which they were derived, *i.e.* the horse-race betting public. The extent to which the results agree with other studies of this type gives some indication of their generality.

Treatment of data. The data have been handled in virtually the same manner as a similar study made by Griffith in 1948.[6] He used data from 1386 races and divided the horses into 11 groups according to the odds established on each horse in the pari-mutuel wagering. By checking the outcome of these races, the true or objective probability (P = winners/entries) was found for each odds-group, and these odds were compared with the subjectively established public odds. In the present study, the total sample has been broken down into eight subsamples depending on the order of the race in the daily program. Each horse whose track odds to win (all are given as odds to one dollar) fell between 0.05 and 25.95 was placed in one of nine groups. The class intervals for the odds-groups were: 0.05–1.95; 2.00–2.95; 3.00–3.95; 4.00–4.95; 5.00–5.95; 6.00–7.95; 8.00–10.95; 11.00–15.95; and 16.00–25.95. Odds of greater than 25.95 were not recorded because the results would not have been sufficiently stable to be of use in the analysis.

The objective probability, P, that a horse in a particular odds-group will win the race is W/N, where W is the number of winning horses in the odds group, and N is the number of entries in that group. In discussing the expected value of bets for the various odds-groups, it is more convenient to use the expectation, E, found from the actual ratio of the amounts of money wagered, *i.e.* from the odds that would have prevailed had not the track take and breakage been deducted.[7] When odds are treated in this manner, positive, zero, and negative values of E have their conventional meaning. Expected value for a $1 bet to win in a particular track odds group is: $E = P \cdot a^* + (1 - P) \cdot - 1$ where a^* = mean corrected winning odds for a particular track odds group.

Reliability of the data. The approximate standard error of E is $\sigma_p (a^* + 1)$ where σ_p is the standard error of P. The standard error of a^* for a particular odds-group is so small compared to σ_p that it can be ignored. Because of the skewness of the sampling distribution for P at the extremes, it is usually not permissible to interpret the standard error of a proportion when P is as small as some of those appearing in this study, *i.e.* 0.05. The size of N in the present case is, however, very large (500–10,000), and under such conditions the standard error of P is applicable as a measure of reliability.

RESULTS

Expected values for entire sample. In Table II and Fig. 1, the expected values are given as functions of odds for the total sample of 9248 races.[8]

---

[6] R. M. Griffith, Odds adjustments by American horse-race bettors, this JOURNAL, 62, 1949, 290–294.

[7] Track odds may be corrected for track take and breakage as follows: $a_c = [(a + 1.025)/(1 - t)] - 1$, where a_c = corrected odds, a = track odds, and t = track take. In the above equation 1.025 represents the original $1.00 bet plus the correction for breakage. With the exception of Fig. 1, expected value, E, always refers to values computed from *corrected* odds. The odds-groups, however, are stated in track odds.

[8] Some examples of the raw data from which the entries in Table II were computed may be helpful to the reader. In the 1156 races which occupied the first position in the day's program, there were 771 horses whose track odds-to-win were 0.05–1 to 1.95–1. Of these, 320 won and returned an average of $1.24 for each

→

TABLE II

EXPECTED VALUES OF ONE DOLLAR BETS AS A FUNCTION OF TRACK ODDS

		Track odds								
Position of Race	Number of Races	0.05– 1.95	2.00– 2.95	3.00– 3.95	4.00– 4.95	5.00– 5.95	6.00– 7.95	8.00– 10.95	11.00– 15.95	16.00– 25.95
1	1156	.08	.04	.05	−.11	−.06	.04	−.10	.12	−.11
2	1156	.14	.13	−.05	.02	.08	−.06	−.21	−.08	−.05
3	1156	.05	.09	.06	.08	−.07	.01	−.12	−.13	−.07
4	1156	.05	.10	.04	.12	−.05	−.12	−.06	−.03	−.10
5	1156	.03	.10	−.02	−.07	.02	−.07	−.02	−.02	−.14
6	1156	.11	.03	−.01	−.05	−.06	−.07	.01	−.05	−.20
7	1156	.01	.00	.00	.08	.03	.19	−.17	−.13	−.32
8	1513	.22	.11	−.09	.04	.09	−.11	−.15	−.21	.03
1–8	9248	.08	.08	−.01	.01	.00	−.03	−.11	−.07	−.11
$(\sigma_E)_{1-7}$		.048	.064	.077	.091	.108	.086	.095	.114	.121
$(\sigma_E)_8$		.053	.056	.057	.069	.089	.073	.071	.082	.121
(σ_E) Total		.017	.022	.026	.032	.038	.031	.033	.039	.044
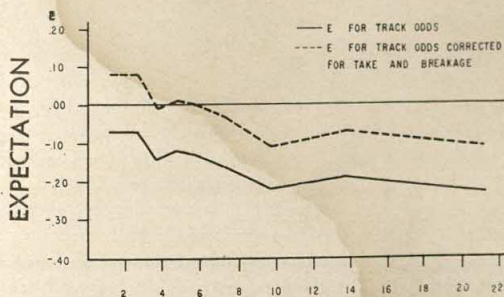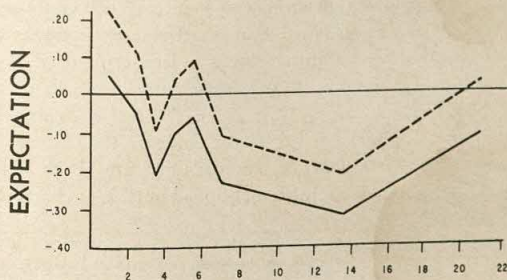
FIG. 1. EXPECTED VALUES OF ONE DOLLAR BETS AS A FUNCTION OF ODDS.

dollar bet, in addition to the original wager. The mean odds δ $(a*)$ that would have obtained had the track take and breakage not been deducted are: [(1.24 + 1.025)/0.87] − 1 = 1.60. From this, $E = 0.415(1.60) + 0.585(−1) = 0.08$. In the total sample of 9248 races over all eight positions in the racing program, there were 8781 horses entered at odds of 8.00–1 to 10.95–1. Of these, 661 won at an average odds of 9.32–1. The expected value is then −0.11.

A horizontal line is drawn on the graph at an expectation, E, of zero. Odds below 3–1 show an E of 0.08, odds of 3–1 to 6–1 have an E approximately equal to zero, and odds greater than 8–1 have an E of about −0.10. The E of 0.08 exceeds 0.00 by four standard errors; the difference between the obtained E of −0.10 and that of 0.00 is significant beyond the 5% level of confidence. The indifference point, where the graph of E is equal to 0.00, is located between odds of 3.5–1 and 5.5–1, or at the probability value of 0.15 to 0.22. This agrees well with Griffith's findings of 0.16 and 0.18 in two similar analyses.[9]

Expected values for subsamples. The primary purpose of this study was to investigate what change, if any, took place in the expected values-vs.-odds relationship as the racing day proceeded. To investigate any trend that might exist, the 9248 races were broken down into eight samples of 1156 races, depending on the position the race occupied in the daily program. Table II gives the expectations (E) for each of the eight races, and standard errors for E of each odds-group. In general, the pattern of positive E for low-odds horses and negative E for the higher odds holds for the subsamples, as it does for the total sample. The first six races all yield E-vs.-odds patterns that do not differ from the pattern for the total sample by more than the sampling error.

The group of seventh races exhibits several interesting features. First, the E for odds of less than 3–1, which has been positive for all the other subsamples, is found to be roughly equal to zero. Secondly, the E for odds of 7–1 is much larger (0.19) than usual, being significantly greater (beyond the 5% level of confidence) than the corresponding value for the total sample. The third feature is the very low E (−0.32) for odds of around 21–1. This is the lowest E for any of the samples. While this pattern is quite different from those of the remainder of the data, the differences are probably not related to betting behavior on previous races. Neither of the adjacent races (sixth and eighth) shows a similar pattern. The explanation of the E-vs.-odds pattern for the seventh race apparently lies in its uniqueness as the feature race of the day. The relatively low E for the low odds may be due to the increased familiarity of the public with two or three of the favorite horses in this race. These horses usually have impressive records and are highly publicized in the local newspapers. Other horses are seldom mentioned outside of charts showing entries and results.

The group of eighth races gives the most interesting results of all the subsamples. The graph of E-vs.-odds for the eighth races, in the lower half of Fig. 1, shows two significant features. First, and most outstanding, is the E of 0.22 for odds below 2–1. This is significantly above 0.00 beyond the 1/10% level of confidence and above the corresponding value for the first seven races (0.07) beyond the 2% level. The second feature of the eighth-race graph is the sharp dip in expectation, E, at odds of 3.5–1. The E is −0.09 compared to 0.01 for the first seven races, being not quite significant at the 5% level of confidence.

In view of the relatively high E for odds of below 2–1 in the eighth races, an effort was made to investigate further this odds-group. As was explained earlier,

[9] Griffith, *op. cit.*, 290-294.

it was not possible to obtain the odds to place and show established on each horse as was the case in the win category. We may, however, categorize the data on the basis of *winning* odds as before, and then list the number of horses that finished first or second, *i.e.* placed, and the number that finished third or better, *i.e.* showed. The place and show odds are available for these horses and can be tabulated as before. The E of a place wager on horses whose odds-to-win were below 2–1 was found to be 0.24, or 6.5 standard errors above 0.00.

Unlike the seventh race, the eighth race is not unique in type. It is almost always very similar in make-up to three or four of the earlier races. The fact that the E-vs.-odds graph for the eighth race is quite different from that for the first seven races must be explained by a change in betting behavior, and this change is due to the position of the race in the daily program rather than the composition of the race. Horses with a high probability of winning, but with accompanying low pay-offs, become even more unpopular with the bettors in the last race.

Amount wagered per person as a function of position in racing meet. During the 1953 Hollywood Park season the average amount bet per person during a racing day was $72.70. The average amount paid to the track in the form of mutuel take then was $9.46 per person exclusive of breakage. These figures represent the mean amount bet. Since the distribution of bets is positively skewed, due to a few large bets, the median is undoubtedly lower than the mean. Probably the former is around $50.00 bet and $6.50 lost. There was a tendency for the amount wagered per person to increase slightly as the seasonal meet proceeded. The average amount bet per person per day for the first 10 days was $67.10 compared to $75.50 for the last 10 days. The weekday average was $76.60 per person as compared to $65.10 per person for Saturdays and holidays.

Relation between amount bet on a race and its position in the daily program. During the racing day the total amount wagered per race ordinarily increases for each succeeding event up to the eighth race. There is usually a slight decline in the total mutuel handle from the seventh to the eighth race. For the 1953 Hollywood Park data the increase from the first to the seventh race is fairly regular, with the amount bet in the latter race being about 1.8 times the amount wagered on the former. Some of this change is due to late arrivals and early departures, but the increase in sizes of wagers is clearly much more than can be accounted for by fluctuations in attendance.

Probability-preferences. Recently, Edwards, has reported several well-designed experiments using college Ss and dealing with probability-preferences in gambling.[10] These studies have held constant the E of bets and determined the preference for different probabilities by means of paired comparisons. Eight bets on a rigged pinball machine were used with probability values of 1/8, 2/8, . . . 8/8. In general, the results of these experiments have shown a definite preference for bets involving the probability of success 4/8, and a definite avoidance of the value 6/8. He found that these preferences were still distinguishable in experiments involving unequal E even though such choices violated the maximization-of-expected-value hypothesis.[11]

[10] Edwards, *op. cit.*, 349-364.
[11] Edwards, Probability-preferences among bets with differing expected values, this JOURNAL, 67, 1954, 56-67.

Furthermore, he found that the above probability-preference remained constant for different levels of expected values, thus demonstrating that the preferences exist independently of the attached utility variable.[12] Finally, Edwards carried out an experiment designed to study 'variance preferences' in gambling.[13] A conservative individual, wishing to minimize the variability of his assets over a series of risk-taking events should choose wagers with small amounts bet and high probability of success. Less conservative individuals may increase the variability of their assets, i.e. gamble on a large win at the expense of risking a large loss, by choosing to wager large amounts at low-probability values. Edwards created special situations in which the best strategy for winning bettors consisted of minimizing the variability of their assets, and the best for losing bettors consisted of maximizing the variability of their assets. The results showed that the same preferences for probabilities that had been found in earlier experiments was still the most important factor in predicting choices. Winning bettors did not change their preferences for low cost or high cost bets as the series of choices proceeded, although it would have been in their interest to do so. Losing bettors did tend to choose high-variability bets when good strategy indicated it; however, choosing bets with high variability was of less importance than the preference for a given range of probability.

The present study presents a measure of preferences for certain probabilities and asset variability in horse-race betting. In pari-mutuel wagering, the bettor may choose among win, place, and show bets. The three pools are independent, such that the amount wagered on a horse in one category has no effect on the odds on the same horse in the other two categories. The amount deducted by the track (13%) is the same for all three pools, although the factor of breakage takes a slightly larger amount from the place and show pools, since there is a higher proportion of redeemable tickets in these categories. As mentioned earlier, the typical ranges of win, place, and show odds are around 2–1 to 50–s, 1–1 to 20–1, and 0.5–1 to 6–1 respectively. Thus, the proportion of the total amount of money wagered in each pool gives a measure of preference for probability ranges among alternatives with roughly equal expected values. This is fairly analogous to Edward's measure of preference for particular probabilities. It should be noted that during a series of races these proportions may change for the group without necessarily effecting a change in the pattern of extended values discussed earlier. Fig. 2 gives the proportion of the total amount bet in the win, place, and show categories as a function of the position of the race in the daily program. The graph for the win pool shows an almost linear increase from 0.49 in the first race to 0.60 in the eighth and last race. The proportions bet in the place and show categories show corresponding decreases.

Risk-taking events and their effect on subsequent betting. While the group of bettors as a whole is always losing money due to the track-take, a proportion of them is winning on any given day. The question of differences in behavior between winners and losers has been raised. A partial answer may be obtained by determining what, if any, relationship exists between the odds that the winning horse pays and the amount of money wagered per person in the following race. If each person is assumed to bet the same amount, the proportion of winning bettors would be:

[12] Edwards, The reliability of probability preferences, this JOURNAL, 67, 1954, 68-95.
[13] Edwards, Variance preferences in gambling, this JOURNAL, 67, 1954, 441-452.

$Q = 0.87/(a_1 + 1.025)$, where Q = the proportion of those persons purchasing win tickets who realize a return; and a_1 = odds-to-win for horse finishing in the first position. Place and show odds were not taken into consideration. Using this measure of the proportion of the population holding successful win tickets, we tested the correlation between these values and the amount bet per person in the following race. The data from the 50-day Hollywood Park racing season were used. Saturdays and holidays were eliminated because it has been shown that these days have a smaller amount bet per person than for weekdays. This left a total of 40 racing days
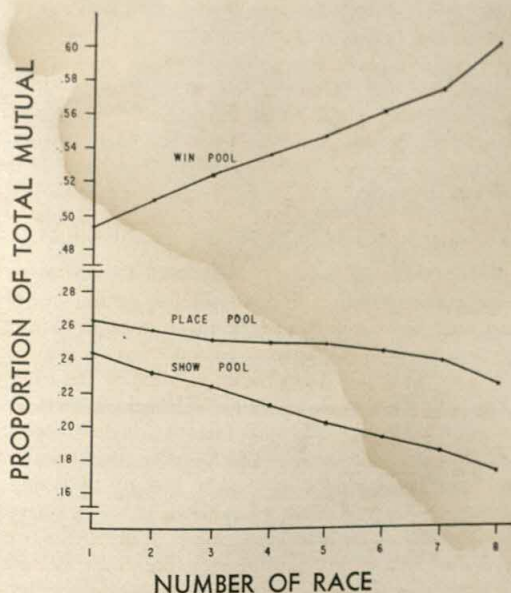
FIG. 2. PROPORTION OF TOTAL MUTUEL BET IN WIN, PLACE, AND
SHOW POOLS AS A FUNCTION OF NUMBER OF RACE
(Data from the 50-day meet at Hollywood Park, 1953.)

and, since there were eight races a day, seven pairs of variables to be correlated. Before computing correlation coefficients, it was necessary to correct the amounts bet per person in each race for the variance contributed by the position of the racing day in the season, since this variable increases as the season proceeds. The seven coefficients ranged from —0.10 to —0.47, with one being significant beyond the 1% level of confidence, and one beyond the 5% level. Since all seven coefficients were negative, they give rather conclusive evidence that bettors increase the amount wagered more after having lost than following a successful bet.

DISCUSSION

Stability of the E-vs.-odds pattern. The general pattern is one of positive expectation, E, for low-odds wagers, high probability) and negative E for

high odds (low probability), with zero E from 3.5–1 to 5.5–1 odds ($P = 0.15$ to 0.22). This pattern appears to have considerable stability inasmuch as it was found with minor variations for the first six subsamples of races. Marked variations of the E-vs.-odds relationship in the seventh race appear to be related to its uniqueness as the feature race. The sharp increase in the E for odds below 2–1 for the eighth race is probably due to certain segments of the population making decisions in accordance with their total financial losses for the day. Bettors apparently refrain from making bets which would not recoup their losses if successful. The increased popularity of odds of around 3.5–1 in the eighth race (indicated by the relatively low E of –0.09) may be due to the fact that a considerable proportion of the population has lost about three times the amount they propose to wager on the last race; however, no evidence was gathered to substantiate this speculation.

In so far as we may generalize to other populations, it appears that subjects can be expected to accept low expected values when low probability-high prize combinations are involved, while demanding higher expected values in the case of high probability-low prize combinations. The central tendency-like effect shows considerable stability over a series of risk-taking events. These findings are consistent with those of Preston and Baratta, who conducted a laboratory experiment on this problem.[14]

Variability of preferences in betting. The betting behavior of the group is such as to increase the variability of their individual assets in an almost linear fashion as the racing day proceeds. This is accomplished by increasing the amount bet per person and by choosing a higher proportion of win-category wagers in preference to place and show betting. This may be partly due to the loss of resources for the group as a whole due to the track take. There is some indication that losing bettors tend to increase the size of their wagers more than do winning bettors.

It is important to note the stability of the E-vs.-odds pattern during the first six races, in spite of the fact that during this same period, size of wagers and preference for low-probability bets (win betting) are steadily increasing. The lack of change in the E-vs.-odds pattern corresponding to an increase in size of wagers would appear to indicate that the utility scale for money in the range considered is virtually the same as the objective dollar scale, and the more important psychological variable is subjective probability. On the other hand, neither is the increasing popularity of win

[14] M. G. Preston and Philip Baratta, An experimental study of the auction-value of an uncertain outcome, this JOURNAL, 61, 1948, 183-193.

betting (low P of success) during the first six races reflected in the E-vs.-odds pattern. If the group's subjective evaluation of low probability-high prize wagers increases over a series of risk-taking decisions, we would expect an intensification of the negative E for high-odds and positive E for low-odds pattern. This does not occur until the last race. These results suggest the inadequacy of a model for predicting risk-taking decisions based solely on the maximization of 'subjectively expected utility.' Allais has suggested that the variances involved in the wager may also be an important factor in this type of decision making.[15] The results of the present study tend to confirm this prediction. This is not in agreement with Edwards' laboratory findings which indicated that variance-preferences were of minor importance compared to probability-preferences in gambling.[16] Edwards also found probability-preferences to be relatively stable, which is not in agreement with the present finding. Perhaps the discrepancy is due in part to the difference between college and horse-race betting populations.

SUMMARY

By means of a statistical analysis of 9605 horse-races, this study sought to obtain information about the stability of decision-making behavior over a series of risk-taking events. In general, the group tended to accept probability-prize combinations whose expected values were less for low-probability wagers than for high ones. This tendency was relatively stable over a series of decisions, and was for the most part, independent of decreasing group resources, size of average wagers, and change in group probability-preferences. The group behaved in a manner such as to increase the variability of their assets as a series of risk-taking events proceeded. This was accomplished by increasing the size of the wager and choosing lower probabilities (win bets) with accompanying prospective higher returns. There was some indication that losing bettors increased the size of their wagers more than did winning bettors.

The relatively stable E-vs.-odds pattern over a series of events in which sizes of wagers and preferences for probability values show consistent changes raises some questions about decision-making models. It was shown that a model making use of subjective probability and utility functions alone does not account for the results found here. In the present study, variance-preferences also play an important role in determining choices among risky alternatives.

[15] Maurice Allais, Le comportement de l'homme rational devant le risque: Critique des postulats et exiomes de l'ecole americaine, *Econometrica*, 21, 1953, 503-546.
[16] Edwards, Variance preferences in gambling, 441-452.

THE CONSTANT-SUM METHOD APPLIED TO
SCALING SUBJECTIVE DIMENSIONS

By Frank J. Dudek and Katherine E. Baker, University of Nebraska

A promising approach to the scaling of psychological variables is the constant-sum method, proposed by Metfessel and developed by Comrey,[1] which requires O to indicate the relative magnitudes of two stimuli by dividing 100 points between them. One advantage of this method is its potential generality of application. Fractionation and multiple-stimulus methods for deriving ratio scales must provide O with a variety of stimuli large enough for him to be able to choose those with specified relations to certain standards. Where stimuli to be scaled have some simple physical property correlated with the psychological dimension judged, this requirement is not too difficult to meet, but otherwise the selection of comparison stimuli is difficult. The constant-sum method, however, does not call for any stimuli over and above those to be scaled. Theoretically, at least, this method may be employed whenever a set of stimuli has a common dimension to be judged with respect to magnitude, and it is equally as easy to use when a corresponding physical measurement exists as when no such measurement is available. To date, investigations of the constant-sum method have involved dimensions for which simple physical scales were available, thus providing useful reference-points to which obtained psychological scales might be related.[2] If agreement among Os is to be found, it is most likely to be evident in such experiments. As yet, however, there is no evidence that Os can make the required ratio-judgments consistently where purely subjective dimensions are concerned.

[1] Milton Metfessel, A proposal for quantitative reporting of comparative judgments, *J. Psychol.*, 24, 1947, 229-235; A. L. Comrey, A proposed method for absolute ratio scaling, *Psychometrika*, 15, 1950, 317-325.

[2] J. P. Guilford, *Psychometric Methods*, 2nd ed., 1954, 208-220; J. P. Guilford and H. F. Dingman, A validation study of ratio-judgment methods, this JOURNAL, 67, 1954, 395-410; Trygg Engen, An evaluation of a method for developing ratio-scales, this JOURNAL, 69, 1956, 92-95; K. E. Baker and F. J. Dudek, Weight scales from ratio judgments and comparisons of existent weight scales, *J. exp. Psychol.*, 50, 1955, 293-308.

In the two experiments reported here, the constant-sum method was employed to scale 'roughness' of sandpapers and 'preferences' for neckties. The aspect of roughness was chosen to represent a case for which the physical correlate is unknown, but for which, at the same time, the concept that a metric exists is not difficult for O to grasp. While physical concepts of friction, work required to move an object over the surface, size and density of granules are relevant to the perception of roughness, the fact remains that the necessary and sufficient physical energy variables are not known to the psychologist, much less to O. The preference for neckties may be considered to represent an even more subjective dimension, for which physical measurement seems quite irrelevant.

EXPERIMENT I

Method. The Os were five graduate students in psychology, all familiar with the constant-sum method. Each O repeated the full set of judgments 5 times in 10 experimental sessions.

Stimulus-materials were nine patches (A–I) of sandpaper ranging from very fine emery paper to very coarse heavy-duty paper.[3] Samples were so selected that each was discriminably different in roughness from every other sample, as demonstrated in preliminary experimentation, and an effort was made to eliminate discriminable features other than roughness, such as a texture-difference, which seemed to be present for emery cloth and garnet paper of about the same roughness.

Nine samples provided 36 different pairs of the papers; with order taken into account, 72 pairings were available for judgment. The pairs were divided into two blocks of 36 trials, one for each of two experimental sessions. Each block included all possible pairs, half in the rougher-smoother order, half in the smoother-rougher order, and the blocks were alike except for reversal of order of presentation of a pair. The order in which pairs were presented was random, and a check showed rough and smooth samples to be quite evenly distributed over the 36 trials.

The papers were presented in a box-like apparatus open only at the rear, thus eliminating visual cues. A small hole in the front side permitted O to reach into the box where the stimulus-patches appeared beneath two 1.5 × 2-in. rectangular apertures placed one behind the other in the floor of the box. The sandpaper patches were glued to two sliding strips of wood which could be moved by E from the rear of the apparatus to place any pair of stimulus-patches beneath the apertures. The duplicate patches on the two sliding mechanisms were cut from the same larger sheet and the same patches were used throughout with the exception of the two finest papers where obvious 'clogging' with use dictated replacement from the larger sheet half-way through the experiment.

Detailed instructions and preliminary practice-trials covering the full range of

---

[3] A through G were from Minnesota Mining & Mfg. Co. types Wetordry Tri-m-ite No. 320 A and 240A (soft back), 3M Imperial Flint No. 1/0, 1/2, Medium, 1½, and 2, respectively. H was Behr-Manning Garnet 36-D2. I was Skil Sander belt, Alum 0X3A from Skilsaw, Inc.

roughness were given at the beginning of the first experimental session. The Os were instructed to divide 100 points between members of each pair of sandpaper patches so as to express relative roughness. They were cautioned against unnecessarily restricting their point-assignments to even 10s or even 5s, and reminded that small variations in point-assignments near 90 imply larger changes in ratios than do comparable variations near 50.

Several restrictions were placed on the exact way in which O was allowed to feel the stimulus-papers. The nearer sandpaper sample always was felt first, then the farther sample, following which O was free to move back and forth between samples until ready to report his judgment. Only one hand was used throughout, and only one finger was used on a given trial. Since continued feeling of sandpapers, some quite rough, could be expected to alter the sensistivity of receptors, the index finger was used on the first trial, the second finger on the second trial, the third finger on the third trial, the index finger again on the fourth trial, and so on. Variability in sensitivity among fingers was taken to be less than the variability produced by prolonged stimulation of the same finger. Movements of the finger were restricted to those in a direction perpendicular to the front wall of the apparatus, since circular and side-to-side movements appeared to produce somewhat different qualitative impressions.

Results. The procedure for deriving scale-values for stimulus-objects from point-assignments has been described elsewhere.[4] Scale-values were computed by averaging the eight estimates of B/A, C/A, D/A, E/A, F/A, G/A, H/A, I/A, instead of by successive multiplication of average ratios between adjacent stimuli in the ranked order as Comrey proposed.[5]

The data of the present experiment provide five independently determined sets of scale-values for the sandpapers for each of five Os, as well as five such sets of scale-values for group-averages. Fig. 1 presents a graphical summary of the findings. For each O, Papers A through I are placed along the psychological scale in positions indicated by the average of scale-values found in the five repetitions of the experiment. Above this line, dots show the five scale-values averaged. A similar treatment for group-data also is included. Here the dots refer to the average of the five individual scale-values on each repetition of the experiment. It should be noted that a logarithmic scale is employed to make equal proportionate differences in scale-value show up as equal.

Of major interest in the present connection are the degrees of agreement exhibited among Os and from one repetition to another for the same O. All Os agree on the rank-order of the sandpapers when the data for the five repetitions are averaged. They also agree on a number of more specific relationships among the stimulus-papers. Papers B and C, for example, are judged to be not very different in roughness by all Os and the propor-

[4] Baker and Dudek, *op. cit.*, 293-308.
[5] Comrey, *op. cit.*, 317-325.

tionate difference in roughness between B and C is nearly the same as between F and G. Furthermore, in all cases the B-C and F-G relationships are judged to be smaller than the E-F and G-H relationships. These statements hold even for the most deviant Os.

Each O was fairly consistent in his five repetitions although a single repetition involved only two judgments of any given pair. In Fig. 1,
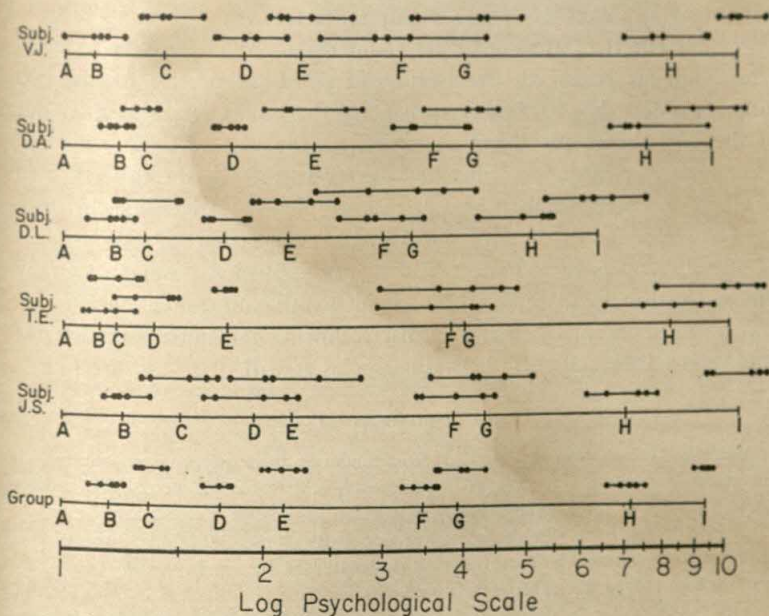
FIG. 1. AVERAGE PLACEMENT OF SANDPAPER SAMPLES A-I ON A PSYCHOLOGICAL SCALE OF ROUGHNESS
(Dots signify scale-values obtained in the five independent repetitions of the experiment.)

overlap in the ranges for adjacent stimuli does not represent reversals in judged order of papers in the different sessions; rather, variations from one repetition to another were in the nature of small expansions and contractions of relationships. Of the 200 possible reversals between adjacent stimulus-papers in five repetitions by five Os, only seven reversals occurred; with two Os showing none. Five of these reversals involved papers B-C or F-G, indicating that these particular combinations are close to the *jnd* for roughness. The reliability of scale-value determinations appears to be about the same throughout the whole set of stimulus-papers.

The fact that the ranges of values for various papers do not differ systematically means that variations represent about the same proportionate (not absolute) changes for all parts of the scale covered.

The group-data are very stable indeed, and scale-values obtained do not differ markedly from one repetition to another. The greater consistency of averaged scale-values does not imply that different Os within the group agree with each other better than individual Os agree with themselves, but that averaging makes the results highly repeatable.

Possible effects of time-order are counterbalanced in the sets of scale-values thus far presented. Examination of point-assignments for the two different orders showed, however, that, for the group averages, in 27 out of the 36 pairings the smoother sample of the pair received a slightly greater number of points when it came first; that is, the difference in roughness between members of a pair of samples appeared somewhat less. Instances in which the opposite was true involved no special combinations of stimulus-papers to indicate that this bias applied to similar pairs any differently than to dissimilar pairs. Evidence that time-order is effective was obtained in spite of the fact that following the initial exposures O could move back and forth between samples at will.

EXPERIMENT II

Method. The Os were volunteers from undergraduate psychology courses. To investigate possible sex-differences in preference, and to have some Os give ratings on two occasions, three independently constituted groups were employed: Group A, 29 men who rated the stimuli only once; Group B, 19 men who judged the stimuli in two sessions a month apart, hereafter designated as B_1 and B_2; and Group C, 29 women who judged the material only once.

The stimulus-objects judged were eight equal-sized, printed, color-reproductions of neckties. To control color as a variable, the samples chosen were predominantly blue. There were 28 possible pairings of the 8 ties, and the order of presentation for judging was random except for the restriction that no tie appeared in successive pairs. Each pair was shown on a large screen by means of an opaque projector. The Os were instructed to divide 100 points between the members of each pair so as to indicate their relative degree of 'liking' for the two ties. These judgments were made following a series of 20 judgments of line-lengths, which provided some experiences in making judgments in terms of point-divisions.

Results. Procedures used to compute scale-values were the same as in Experiment I. Scale-values of the neckties as determined for the various groups are shown in Fig. 2.

It is apparent from Fig. 2 that the two groups of men agree quite closely with respect to the order of stimuli. One index of this agreement

is *tau,* a coefficient of agreement based on the number of interchanges necessary to make two sets of orders correspond. With two interchanges the order of items for Group A corresponds to the order for Group B_1 ($\tau = 0.86$); and only one interchange results in orders identical for Group A with those from Group B_2 ($\tau = 0.93$). Three interchanges are required to make orders for Group B_1 and B_2 comparable ($\tau = 0.79$). Thus it appears that two independent groups of judges agree as well as do two judgments made by the same group on separate occasions. A mere count of interchanges does not reflect certain important uniformities mani-
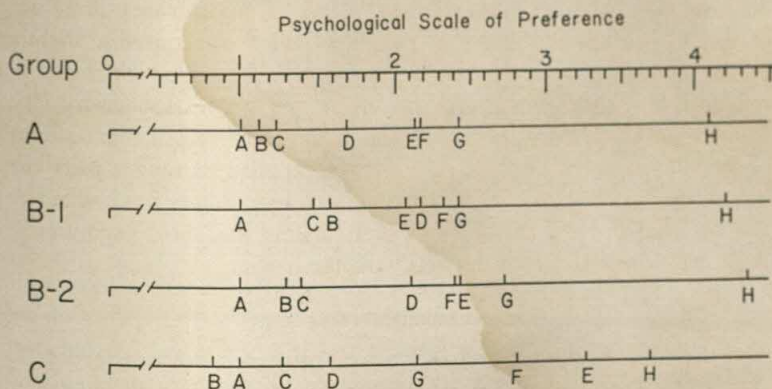
FIG. 2. SCALE-VALUES OF EIGHT NECKTIES AS DETERMINED FOR THREE GROUPS OF JUDGES

fested in the sets of scale-values for the groups of men. Analogous with the roughness judgments, rather striking consistencies with respect to the relative distances maintained between stimulus-objects are evident from group to group.

In general, the agreement between orders for women and men is somewhat poorer. The 4, 6, and 3 interchanges required to bring about corresponding orders yield τs of 0.71, 0.57, and 0.79, respectively. Although all coefficients are significant, sex-differences with respect to preference for the neckties are suggested.

DISCUSSION

One aim of these experiments was to determine if *O*s could make ratio-judgments of the sort required by the constant-sum method when the

stimulus-objects scaled had no simple physical correlate. The answer to this question depends upon the degree and type of consistency demanded. It seems reasonable, however, to conclude that two psychological metrics have been demonstrated, since consistencies of judgment belie any interpretation of the point-assignments as merely random guesses.

Lack of complete agreement in the data does not contradict the claim that a metric exists, since agreement among individuals or among groups is not crucial, in and of itself, to validating the methodology as such, or in specifying the properties of obtained scales. While an individual O should be consistent with himself, there is no reason why two individuals or groups of judges should necessarily agree in evaluations of subjective material. With physical dimensions like length and weight, perceived relations among stimulus-objects may be expected to remain more or less constant. For subjective dimensions, even though Os may agree on the meaning of "feels rougher than" or "like four times as well," they need not agree on which item should get a particular scale-value. For preferences especially it might be expected that specific item-values would change over a period of time although the metric did not vary.

If we accept the conclusion that psychological scales have been demonstrated, a further question concerns the nature of the scales obtained. In contrast to simple ranking, this method gives numerical scale-values which imply ratio-relationships, and it is important to know to what extent we should place confidence in the exact numerical values. For roughness judgments, there appears to be no doubt that a consistent rank-ordering of stimuli results from application of the constant-sum method. Furthermore, in spite of variations in scale-values, certain uniform relations among stimulus-objects within our set were shown to be common to all Os. Variations were smaller within than between Os, perhaps signifying individual differences in the metric itself. Preferences for neckties is even more likely to be an individual matter. Thus, to find agreement between group-scales is reassuring since it then becomes possible to talk meaningfully in terms of scales derived from average values determined for groups. If there were no communality within groups, each member might have his own scale, but no scale would have any generality. With respect to the preference-dimension, also, there were noted consistent distance-relationships between stimulus-objects that would not be expected if only ordinal characteristics were reflected in the scale.

While it might be not only premature but also incorrect to claim that a true ratio-scale has been demonstrated, Comrey's position that there may

be scales whose properties place them between ordinal and ratio-scales would seem reasonable in the present context.[6] If we imagine a ratio-scale, like physical measurement of length, into which increasing amounts of error are introduced, somewhere along the process we do not have our original confidence in the exact scale-values given to objects, but at the same time the scale-values may be taken to have more meaning than mere rank-orders for objects. Even if the constant-sum method achieves only an approximation to a scale with ratio-properties, such a scale seems more meaningful than those which could be derived by alternate methods of ranking or even paired comparisons with certain assumptions about dis-criminal dispersions. The constant-sum method not only asks for more information about relative differences from O but retains this information in the scale-values determined.

From the standpoint of information asked of O, it might be argued that the constant-sum method gives a ratio-scale by definition. A unit of measurement is defined, and a zero-point for the scale is thus implied. Further, relations among scale-values of stimulus-objects remain invariant no matter what stimulus-object is considered as the unit of the scale. As Michels and Helson point out, such scales have genuine operational mean-ing since they are established by a clearly defined and reproducible experi-mental procedure.[7] They raise the question, however, which in a sense seems the crux of the issue, whether "this operational basis is broad enough to allow scales to have psychological meaning in situations other than that involved in their determinations." This criterion would be met when any independent referent (like the sale of neckties, for example) was shown to have a one-to-one correspondence with the scale-values of items, since then clearly the scales would have a psychological meaning, in the ratio-sense, beyond the context within which the scale-values were determined.

Only empirical evidence will reveal whether the degree of variation in scale-values or discrepancies from prediction are tolerable. The constant-sum method appears to be a valuable one, since even with considerable experimental error it is likely that predictions from such determination would be better than those that might be made from existent methods where it is quite certain that no more than ordinal properties can be achieved.

---

[6] Comrey, An operational approach to some problems in psychological measure-ment, *Psychol. Rev.*, 57, 1950, 217-228.
[7] W. C. Michels and Harry Helson, A reconciliation of the *VEG* scale with Fech-ner's law, this JOURNAL 67, 1954, 677-683.

Summary

The constant-sum method was employed to obtain scales for subjective dimensions represented by 'roughness' of sandpapers and 'preference' for neckties. With respect to roughness it was found that Os agreed with themselves and with each other at least on the rank-order of stimuli, and a number of specific relationships within the sets of scale-values were common to all Os. For the preference-dimension, there was a fair degree of agreement between scales determined for different groups and between scales determined for the same group on two occasions. There were, however, sex-differences with respect to the placement of items on this dimension. It was concluded that consistencies in scale-values indicated measurement on scales that reflected more than merely the ordinal characteristics of items, although the requirements of a true ratio-scale may not have been met fully. The implications of these results were discussed and various problems arising in attempting to determine the characteristics of scales determined for subjective dimensions were considered.

AN EXPERIMENTAL ANALYSIS OF SUBCEPTION

By CHARLES W. ERIKSEN, The Johns Hopkins University

Several years ago, Lazarus and McCleary introduced the concept of *subception*, which they defined as "a process by which some kind of discrimination is made when the subject is unable to make a correct conscious discrimination."[1] They invoked this concept to account for the results of an experiment on tachistoscopic recognition: they conditioned the GSR (elicited by shock) to one group of nonsense-syllables but not to a control group, and found that the tachistoscopic exposure of conditioned syllables resulted in a larger average GSR than the exposure of control syllables, even when the duration of exposure was too small for correct verbal identification. Although Lazarus and McCleary attributed the difference to unconscious discrimination, the remainder of this paper will refer to their finding as the 'subception-effect' without any implication of unconscious processes—in keeping with a distinction previously made by Howes.[2] It has been shown elsewhere that the subception-effect can be stated in terms of a partial-correlation model.[3] In the subception experiment, S is required to make two concurrent responses to a stimulus, a verbal response and a GSR. The subception-effect consists of a significant partial correlation between the GSR and the stimulus when the verbal response is held constant.[4]

Formulation of the problem in these terms has several advantages. It places the subception-effect in a broader context and suggests several methods of investigating the underlying variables. One advantage is that the general conditions that must be met for a subception-effect to be obtained can be specified. These are that both responses must be correlated with the stimulus, but not perfectly, and that the two responses must not be correlated perfectly with each other—or, stated somewhat differently, both responses must have a certain amount of error, and the error in

[1] R. S. Lazarus and R. A. McCleary, Autonomic discrimination without awareness: A study of subception, *Psychol. Rev.*, 58, 1951, 113-122.
[2] Davis Howes, A statistical theory of subception, *ibid.*, 61, 1954, 98-110.
[3] C. W. Eriksen, Subception: Fact or artifact? *ibid.*, 63, 1956, 74-80.
[4] The term 'correlation' is used in the general sense and does not imply any particular statistic. The formulation used here applies equally well to non-linear or even non-metric data.

the two response-systems must not be perfectly correlated. These conditions are necessary and sufficient for a partial correlation between one response-system and the stimulus, with the other response held constant (or for the subception-effect). A second advantage of the partial-correlation formulation is that it allows us to focus on the variables that may account for non-correlated errors between verbal and autonomic responses. The awareness of such non-correlated errors between two responses or response-systems is not new to psychology. Hull has recognized them in his concept of asynchrony of behavior oscillation.[5] Several variables that might account for non-correlated errors between verbal and autonomic responses have been discussed elsewhere.[6]

It was the primary purpose of the present experiment to test the partial-correlation formulation of the subception-effect. While Lazarus and Mc-Cleary dealt with a fairly complex perceptual situation, the present formulation of the problem suggests that the subception-effect be obtained in a simple stimulus-generalization task. It was the second purpose of this study to investigate a response-variable that might contribute to non-correlated errors between verbal and autonomic responses.

METHOD

Design. One group of Ss had a *GSR* and a verbal response (the number '6') conditioned to a 35-mm. square stimulus. An electric shock was used as the unconditioned stimulus (*US*) for the *GSR*. After conditioning, these Ss were presented with a series of 11 squares of different sizes—the conditioned square along with 5 larger and 5 smaller squares. The squares were presented one at a time under conditions of absolute judgment. The Ss were instructed to call the smallest square '1,' the largest square '11,' and to assign the intermediate numbers to the intermediate squares in order of increasing size. They were informed that the conditioned square, which they had been trained to call '6,' was actually the mid-square in the series and occupied position 6. They were told that they would be shocked only when size 6 was presented, although that square would not always be accompanied by shock. Each of the 11 squares was presented three times, *GSR*s and verbal responses recorded.

This particular experimental task was chosen since it presents a simplified situation for observing the subception-effect and also fulfills the conditions specified above under which the subception-effect should be obtained. The discriminative problem presented had been found in previous experiments to be well beyond the capacity for perfect performance.[7] In other words, the correlation between the stimuli and the verbal responses would be less than perfect. Furthermore, studies of *GSR*-generalization had shown that differential magnitudes of *GSR* could be expected to the different-size stimuli, and there again the correlation of the response with the stimulus

---

[5] C. L. Hull, *Principles of Behavior,* 1943.

[6] Eriksen, *op. cit.,* 74-80.

[7] C. W. Eriksen and Henry Wechsler, Some effects of experimentally induced anxiety upon discrimination behavior, *J. abnorm. soc. Psychol.,* 51, 1955, 458-463. C. W. Eriksen and H. W. Hake, Absolute judgments as a function of stimulus range and number of stimulus and response categories, *J. exp. Psychol.,* 49, 1955, 323-332.

would be less than perfect.[8] The remaining condition, that the two response-systems have a non-correlated error-term, had already been demonstrated by Lazarus and McCleary.[9] The one thing their experiment did was to show a degree of independence between the GSR and the verbal response. With these conditions met, it was expected that the GSRs of the present study would tend to vary systematically in magnitude with the stimulus when the verbal responses were held constant.

To investigate one source of non-correlated error between the GSR and verbal responses, a second group of Ss was used. It had been found in a preliminary experiment that the number of generalized verbal responses in a discriminative task was a function of the number of specific verbal labels available to the Ss for the non-conditioned stimuli. Those Ss who had only two verbal responses in terms of which to differentiate the conditioned from the non-conditioned stimuli gave significantly more generalized responses than Ss who had a specific verbal response for each of the non-conditioned stimuli.[10] Since this preliminary experiment demonstrated that the number of available verbal responses influenced the verbal judgment of the CS, it was desired to determine whether this variable would also affect the GSR-generalization.

The Ss in the second group received identical treatment with those of the first group during the training phase of the experiment. Prior to beginning the test for generalization, however, the Ss in the second group were instructed somewhat differently—to respond with '6' when the training square occurred and to say 'no' if the square shown was larger or smaller than the training square. (In all other respects the instructions were the same as those described above.) In the remainder of the paper, this group will be designated the 'two-response group,' and the first group the 'eleven-response group.'

Procedure. During the training phase, both groups responded with the number '6' when a black, 35-mm. square stimulus appeared. Conditioning of the GSR was established by a partial-reinforcement schedule (80%) with a moderate electric shock to the left ankle as the US. During the conditioning trials, a blank stimulus was randomly interspersed with the CS on an average of once in every five presentations. All stimuli were exposed for a duration of 3 sec. Shock was administered 1.5 sec. after the onset of the CS and Ss were instructed to withhold their verbal responses until the cessation of the stimulus. Conditioning trials were continued until the S showed a stable, clear-cut GSR on presentation of the CS with but little GSR to the blank stimulus. Following the conditioning trials, the Ss were randomly assigned either to the 'two-' or to the 'eleven-response' group.

Prior to beginning the generalization trials, the Ss in both groups were shown the largest and the smallest squares. Following a reinforced presentation of the CS, the generalization-trials began. During these trials each of the 11 squares occurred once within the first 11 trials in a random order with the restriction that no two adjacent sizes occurred successively. Following these first 11 trials without interrup-

[8] C. I. Hovland, The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone, *J. gen. Psychol.*, 17, 1937, 125-148.

[9] Lazarus and McCleary, *op. cit.*, 113-122.

[10] In the preliminary experiment, the same methods and procedures were employed as in the present study with the excption that GSRs were not conditioned, nor were electric shocks involved.

tion, a reinforced presentation of the *CS* occurred, and then there was another random series of the 11 stimuli. Again another reinforced *CS* was given and followed by the last presentation of each of the 11 stimuli. Care was taken to balance the random order of the presentation of the stimuli as much as possible, and the same random orders were used for both groups. During the generalization-trials, verbal responses and *GSRs*, to each stimulus were recorded.

Materials. The squares ranged in size from 20 to 50 mm. square and increased in size by steps of 3 mm. per side. They were made of black paper pasted in the center of a silver disk 10 in. in diameter. They were presented automatically by a converted record changer. The blank stimulus, used only during conditioning, consisted of a disk without a square, and it was presented with the same noises and procedure that accompanied presentation of the conditioned stimulus. The method and equipment used in presenting both the training- and generalization-stimuli have been described in more detail elsewhere.[11]

Electric shocks were administered through an inductorium activated by a 3-v. source. The intensity of the shock was adjusted prior to the experiment to a level described by the particular *S* as "moderately painful." *GSRs* were continuously recorded by the equipment and circuit previously described by Haggard and Gerbrands.[12]

Subjects. The *Ss* were undergraduate students (men) who had volunteered to serve for pay in psychological experiments. A total of 30 *Ss* was used, of which 10 were assigned to the 'eleven-response' group and 20 to the 'two-response' group.

RESULTS

In Fig. 1, the average *GSR* (in micromhos) elicited by the different squares is shown for the two groups of *Ss*. The general results were consistent with what would be expected from previous work on stimulus-generalization of the *GSR*.[13] The highest average *GSR* was obtained for the 35-mm. square, *CS*, and progressively smaller *GSRs* were obtained for smaller or larger stimuli. The 'two-response' group showed a somewhat sharper response gradient than did the 'eleven-response' group. A modified three-way analysis of variance (stimuli, groups, and individuals) yielded a significant effect only for stimuli ($p < 0.001$). Differences between the two groups did not approach significance ($p > 0.20$) nor did the interaction between groups and stimuli ($p > 0.20$).

While there is no evidence of a difference in *GSR*-generalization between the two groups, results of the preliminary experiment described above suggested that there would be differences between these groups in verbal-response generalization. That such differences occurred is indicated in Fig. 2, each point in which represents the mean number of '6'-responses

[11] Eriksen and Wechsler, *op. cit.*, 458-463.
[12] E. A. Haggard and Ralph Gerbrands, An apparatus for measurement of continuous changes in palmar skin resistance, *J. exp. Psychol.*, 37, 1947, 92-98.
[13] Hovland, *op. cit.*, 125-148.

elicited by the different stimuli for the two groups. As can be seen, rather typical generalization-gradients were again obtained, but with the gradient for the 'two-response' group consistently above the gradient for the 'eleven-response' group. Except for the end-stimuli, which were never mistaken for the CS, the 'two-response' group gave a greater number of '6'-responses to all stimuli including the CS, a finding in agreement with
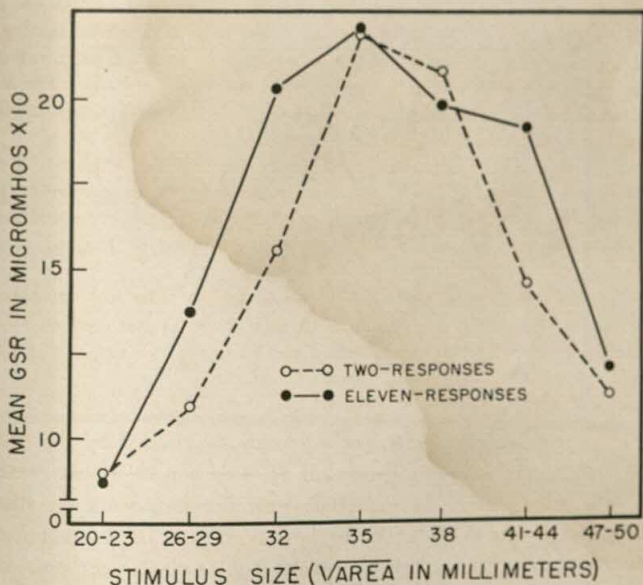
FIG. 1. AVERAGE MAGNITUDE OF *GSR* TO THE DIFFERENT SQUARES IN THE TWO-AND ELEVEN-RESPONSE GROUPS

that of the preliminary experiment. The statistical reliability of this difference between the groups was evaluated by the modified three-way analysis of variance. The data for this analysis were the number of '6'-responses made by each S to each stimulus. To minimize the number of zero-entries in the analysis, the two stimuli at each end of the size-continuum were excluded and the remaining stimuli were combined as shown in Fig. 2. The data were then normalized by the Freeman-Tukey arc-sign transformation.[14] Two significant effects were revealed in this analysis. The first was due to the variation among the stimuli ($p < 0.001$)

---

[14] M. F. Freeman and J. W. Tukey, Transformations related to the angular and the square root, *Ann. Math. Stat.*, 21, 1950, 607-611.

and the second was due to the differences between groups ($p < 0.05$). The interaction between stimuli and groups did not approach significance ($p > 0.20$), which suggests that while the two groups differed with respect to number of generalized responses, the slope of the generalization-gradients did not differ.

The results show that the number of available verbal responses is a significant determiner of verbal-response generalization but has no effect
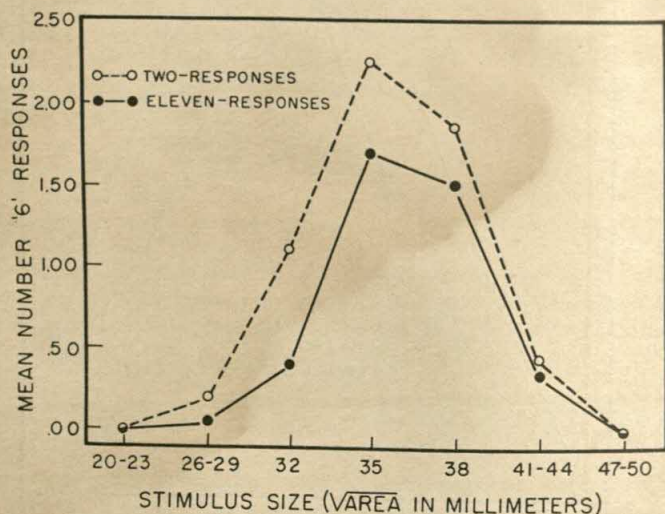
FIG. 2. AVERAGE NUMBER OF 6-RESPONSES TO THE DIFFERENT STIMULI IN THE TWO- AND ELEVEN-RESPONSE GROUPS

upon the generalization of the GSR. Not only is there a lack of significant difference in GSR-generalization between the 'two-' and the 'eleven-response' groups, but the difference which does exist is in the direction opposite from that which would be expected on the basis of the effects due to number of verbal responses on the verbal-generalization gradient. It would seem, therefore, that we have suggestive evidence of a variable that will differentially affect the verbal and autonomic systems.

The formulation of the subception-effect as a partial correlation requires that a correlation be shown between the stimuli and the GSR when the verbal response is held constant. That such a partial correlation does appear in the present data is indicated by Fig. 3, in which the GSR has been separately plotted for all trials on which the verbal response of '6' occurred and for all trials that resulted in a verbal response other than

'6.'[15] The data in the lower curve show the variation in *GSR* as a function of stimuli when the verbal responses were all the same, either '*no*' or a number other than '6.' The upper curve shows the variation in *GSR* as a function of stimuli when all the verbal responses were '6.' The curves of Fig. 3 correspond to the subception-effect of Lazarus and McCleary. The lower curve shows that when the verbal responses were other than '6,' the
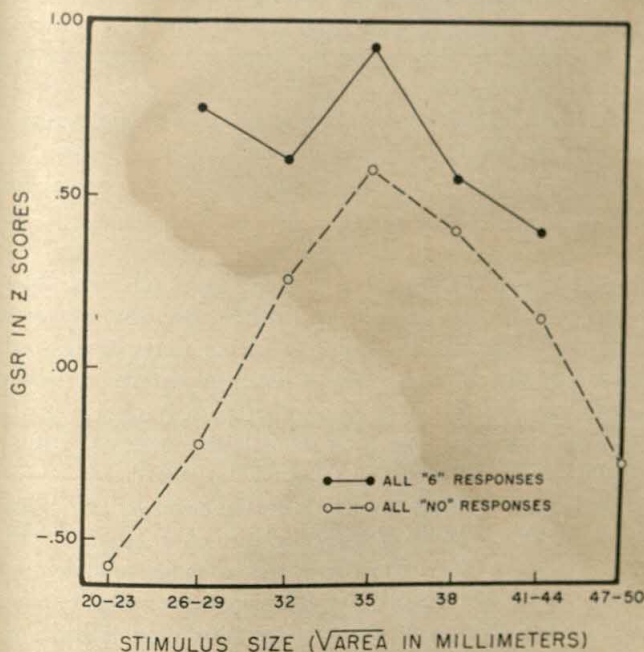
FIG. 3. AVERAGE *GSR* (z-SCORES) TO THE DIFFERENT STIMULI WITH VERBAL RESPONSE CONSTANT

GSR still varied systematically with the magnitude of the stimulus. On the average, the greatest *GSR* was obtained when a negative response was given to the 35-mm. square, the *CS*, while negative responses to larger

[15] The values in Fig. 3 were obtained by converting each *S*'s *GSR* to each stimulus into a z-score derived from the individual *S*'s distribution of *GSRs* in micromhos. By this technique it was possible largely to overcome the limitation imposed by the fact that each *S* did not contribute an equal number of responses to each stimulus in this comparison. The data for the two- and the eleven-response groups were combined due to the small number of cases that occurred for certain of the stimulus-categories. Examination of the data for the two groups prior to combining them had shown no evidence of systematic differences.

or smaller stimuli were of lesser magnitude. In other words, a negative response to the *CS* was accompanied by a greater *GSR* than was a negative response to a non-conditioned stimulus. The upper curve in Fig. 3 illustrates the remainder of the subception-effect. It shows variation of the *GSR* with the stimuli when all the verbal responses were '6.' As can be seen, the *GSR* accompanying a '6'-response to the *CS* was greater than a '6'-response to a larger or smaller stimulus. In other words, the *GSR* was not as great when the '6'-response was used incorrectly as when it was correctly applied.[16] The data in Fig. 3 comprise the subception-effect of Lazarus and McCleary. They show that in the absence of a correct verbal response, the *GSR* on the average will indicate some discrimination among stimuli, or, in the terms of the present conception, that a partial correlation exists between the *GSR* and the stimuli when the verbal response is held constant.[17]

DISCUSSION

The above results confirm the prediction that a subception-like effect can be obtained on a relatively simple discriminative task. Essentially, the only requirement for this effect is that two separate responses be made to the stimulus and that these responses be less than perfectly correlated with each other. Viewed from the standpoint of partial correlation, the subception-effect is not surprising, but is actually the expected outcome. The real experimental problem lies in the variables that determine the non-correlated errors between the two responses. The present study has suggested one of these sources of non-correlated error between verbal and autonomic responses. We have seen that the number of different verbal responses available had a significant effect upon the verbal-response generalization, but there was no evidence that this variable had any effect upon the *GSR*-generalization. While this is suggestive evidence of the kind of variable that can be expected differentially to affect verbal and autonomic responses, it probably plays but little role in the subception-effect obtained here.

The subception-effect in the present experiment can be directly attributed to a fundamental difference between verbal responses and *GSR*s. The verbal response is a discrete response while the *GSR* is for all prac-

---

[16] The reversal in the upper curve for the 26- and 29-mm. stimuli probably is a chance fluctuation, since this point is based upon only three scores, two of which are from the same *S*. All other points in the figure are based upon five or more scores.

[17] Due to the confounding of *S*s with stimuli that necessarily occurs in this analysis, there is no ready statistical technique for evaluating the significance of this subception-effect, but several approximate techniques were used, all of which gave significance values beyond the 5% level.

tical purposes continuously distributed. The Ss in the present experiment could say 'yes' or 'no', or give one of 11 numbers, but the possibilities for verbal response were definitely limited, while for all practical purposes their GSRs could assume infinitely many values. If a S was confronted with a given square and was uncertain as to whether to call it a '6' or '7,' he was forced to give one or the other of the two verbal responses, but the GSR, because of its more nearly continuous distribution, was capable of a compromise response. Thus we have a situation in which the GSR gives information beyond that obtained from the verbal response.

What bearing do the results of the present study have upon the problem of unconscious discrimination? The answer depends largely upon definition. Lazarus and McCleary have implicitly equated conscious discrimination with the presence of a discriminated verbal response. There are several faults to be found with such a definition, but even if we disregard them and accept the definition, neither the present experiment nor that of Lazarus and McCleary permits conclusions in terms of unconscious discrimination. It is not sufficient to assume that if a discriminated verbal response is absent, no conscious discrimination is taking place, or that S is incapable of making a discriminated verbal response to the stimulus. Before conclusions can be drawn about unconscious discrimination, it is necessary that the experimental design be such as to show that the S had available for use a sufficient number of verbal responses to identify separately each discrimination he was capable of making. Neither the present experiment nor that of Lazarus and McCleary meets this requirement.

In the present study, if we consider the discriminations made by the Ss of the 'two-response' group, we find that they always said 'no' to the 20-mm. stimulus and nearly always to the 29-mm. stimulus. On the basis of the verbal response there is no evidence that these Ss can tell the difference or discriminate between these two squares, but it would obviously be fallacious to conclude that because the GSR shows discrimination, evidence of unconscious discrimination has been found. The experimental arrangements simply did not provide the Ss in this group with any verbal responses that they could use to reflect discriminations they were capable of making between the two squares. The data for the 'eleven-response' group shows that Ss were capable of 'conscious' discrimination when they were allowed the necessary verbal categories.

It has been shown elsewhere that just such an artificial response-limitation existed in the subception-experiment of Lazarus and McCleary.[18]

---

[18] Eriksen, *op. cit.*, 74-80.

Even in the 'eleven-response' group of the present study there was an artificial or experimentally imposed limitation of response. The subception-effect, or partial correlation between the GSR and the stimuli, obtained in the present experiment might well have disappeared if the Ss were also allowed to increase the number of different verbal responses by using second guesses to reflect their judgmental uncertainties. The results obtained by Bricker and Chapanis would indicate such an outcome.[19]

SUMMARY

The subception-effect of Lazarus and McCleary was analyzed as a problem in partial correlation, where the effect may be defined as a significant partial correlation between the stimuli and the GSRs when the verbal responses are held constant. It was shown that the necessary and sufficient conditions for such an effect would be satisfied if both the verbal responses and the GSRs are correlated with the stimuli, but less than perfectly, and if each of these two response-systems contained independent error-terms. To test the adequacy of this formulation, which suggested that the subception-effect should be obtainable under a wide range of conditions, an experiment on stimulus-generalization was performed. Two groups of Ss were conditioned to give concurrently a GSR and a verbal response to a square of given size. During the generalization-trials, one group of Ss was allowed to use 11 different verbal responses and the remaining group was permitted only 2 verbal responses. The Ss with only two verbal categories were found to give significantly more generalized verbal responses, but there was no evidence of any difference in GSR-generalization between the two groups. The data from both groups showed evidence of the subception-effect. When verbal responses were held constant, the magnitude of GSR varied systematically with the size of the stimulus. The question of unconscious discrimination was considered and it was pointed out that neither the present experiment nor that of Lazarus and McCleary contained the necessary control operations to permit conclusions or interpretations in terms of unconscious discrimination.

---

[19] P. D. Bricker and Alphonse Chapanis, Do incorrectly perceived tachistoscopic stimuli convey some information?, *Psychol. Rev.*, 60, 1953, 181-188.

FIGURAL AFTER-EFFECTS WITH TACHISTOSCOPIC PRESENTATION

By ALLEN PARDUCCI, Swarthmore College, and KENNETH BROOKSHIRE, University of Oregon

Several investigators have reported failure to obtain figural after-effects when the inspection-figure (*I*-figure) was exposed for less than 5 sec.[1] Köhler and Wallach obtained the usual after-effects, however, by presenting either 10 half-second or 20 quarter-second exposures of the *I*-figure before presenting the test-figure (*T*-figure). They also reported after-effects for some *O*s after a single *I*-exposure as short as 1 sec., providing there was no interval between *I*- and *T*-exposures.[2] No attempt to control the duration of the *T*-exposure was reported for any of these experiments; and since *O*s were not forced to make immediate judgments,[3] the effects of *I*-exposure may have completely disappeared during the course of the *T*-exposure. This possibility is in line with the results of a recent experiment by Krauskopf, which demonstrated that the after-effect may decrease significantly during the first 1.5 sec. of the *T*-exposure.[4]

The present experiment explores the possibility of obtaining after-effects following tachistoscopic presentation of the *I*-figures when the exposures of the *T*-figures also are brief. The hypothesis is that after-effects increase with the duration of the *I*-exposure and decrease with the duration of the *T*-exposure; after-effects are expected when both *I*- and *T*-exposures are brief.

An additional variable to be explored in this experiment is the length of the interval between *I*- and *T*-exposures. In previous experiments, in which the tendency toward after-effect was found to be a decreasing function of the *I-T* interval, short values of this interval (*i.e.* values within the range of the exposures used in the present experiment) were not investigated.[5] In addition to using shorter intervals, a particular effort is made here to prevent exposure to background stimuli which might tend to weaken the expected after-effects. Under these conditions, the importance of the interval between *I*- and *T*-exposures may be reduced sufficiently to permit the appearance of after-effects with tachistoscopic *I*-exposure.

METHOD

Observers. Forty-two volunteers from a class in introductory psychology were assigned at random to the various experimental conditions.

---

[1] J. J. Gibson and Minnie Radner, Adaptation, after-effect and contrast in the perception of tilted lines: I. Quantitative studies, *J. exp. Psychol.*, 20, 1937, 453-467; Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. Philos. Soc.*, 88, 1944, 269-357.

[2] Köhler and Wallach, *op. cit.*, 355.

[3] Elaine Hammer (Temporal factors in figural after-effects, this JOURNAL, 62, 1949, 337-354) reported that the method of adjustment, used by herself, and by Gibson and Radner, typically involves *T*-exposures of from 6-8 sec.; and the latter *Es* set a limit of 8 sec.

[4] John Krauskopf, The magnitude of figural after-effects as a function of the duration of the test-period, this JOURNAL, 67, 1954, 684-690.

[5] Hammer, *op. cit.*, 337-354; J. F. Bales and G. L. Follansbee, The after-effect of the perception of curved lines. *J. exp. Psychol.*, 18, 1935, 499-503.

Materials. Stimulus-materials of the type illustrated in Fig. 1 were drawn with India ink on square white cards. The maximal dimension of the figures varied from 0.25-2.5 in. There were 12 different pairs of *I*- and *T*-figures, 2 warm-up figures, and 15 extra figures randomly inserted between some of the pairs to break up guessing habits (*i.e.* a possible tendency to alternate responses).[6] The usual fixation-points were omitted from all figures in order to reduce possible normalizing effects which any stimulus-elements common to both *I*- and *T*-figures might be expected to produce. Preëxperimental exploration suggested that after-effects could be obtained without fixation-points when the exposure-periods were brief.

The stimulus-materials were exposed, one at a time, in a Gerbrands Harvard Tachistoscope. The 'preëxposure' side of the tachistoscope contained at all times a plain white card with a single, centrally located fixation-point (to orient *O* toward

I - A I - B I - C

T - A T - B T - C

a　　　　　　b　　　　　　c
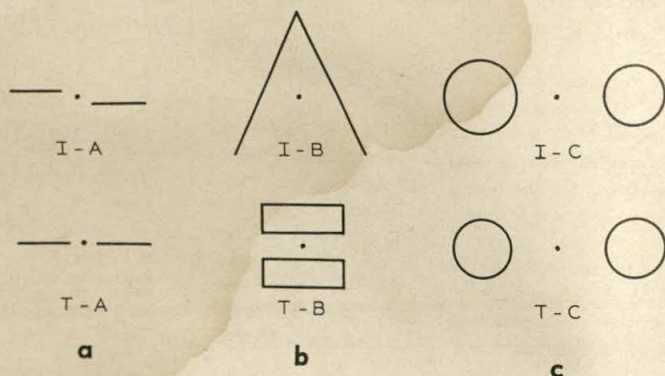
FIG. 1. SAMPLE *I*- AND *T*-FIGURES

the center of the subsequent *I*- and *T*-figures and away from the circular border of the visual field). This is what *O* saw whenever the stimulus-figures were not being exposed.

To reduce the effects of background-stimuli, the testing room was kept dark throughout the experiment. Exposure to the stimulus-cards was monocular, through a single round reduction-screen (the round border would presumably have less of an anchoring effect upon *O*'s judgments for most of the judged dimensions than the otherwise rectangular border of the exposure-apparatus). The angle subtended by the entire visual field was 11°26′ and by the largest figure, 5°44′.

Procedure. The following instructions were read to all *O*s at the beginning of the experiment:

In this experiment you will be asked to discriminate the relative sizes and shapes of objects when they are presented to you for short periods of time. Some of the discriminations will be easy to make, others not so easy. Previous research has shown, however, that people can make correct judgments of things even when they do not think they can; so always try your best; you will be right more often than not.

---

[6] The use of successive pairs of figures may produce inter-pair effects. An effort was made to select pairs which would not systematically enhance the after-effects obtained with subsequent pairs.

In this experiment, I wish you to tell me which figure—the one on the left or the one on the right—is larger than the other. Please do not make your judgment until the test-card has disappeared—that is, until the dot which you now see returns to your vision. Also, between exposures, please do not remove your head from the head-rest—keep looking at the dot; for I will continue flashing stimulus-cards, and I don't wish you to miss any of them. Tell me, then, which of these figures is the larger?

Before the presentation of every stimulus-figure, O was told what kind of judgment he should make (relative size, length, height, separation, and the like), which was determined by the expected distortion of the T-figures. Thus, for Fig. 1a, O was told to report, both for the I- and T-figures, which of the two lines appeared lower. For Fig. 1b, he was to report nothing following exposure of I; but following exposure of T, he was to report which of the two rectangles was wider. For Fig. 1c, he reported which circle was larger in both I- and T-figures. Four of the I-T pairs involved two judgments while the remaining eight involved only one.

Three different exposure-periods were used: 0.25, 0.75, and 5.0 sec. For any given O, all I-figures were exposed for one of these periods; all T-figures were also exposed for one of these periods, though not necessarily for the same period as the I-figures. The interval between I- and T-exposures was 3 sec. for half the Os and 10 sec. for the rest. The experiment thus involved the manipulation of three variables in a $3 \times 3 \times 2$ factorial design. Assignment to the respective experimental conditions was random, two Os being run under each condition.

Six additional Os were used as controls. These Os were exposed to the 12 T-figures only; they were run through this series twice, two Os being exposed for each of the three exposure-periods. Their judgments thus provided a baseline against which the significance of obtained after-effects could be assessed.

RESULTS

A single measure of after-effect was computed for each O. Judgments in the expected direction (e.g. left T-line lower in Fig. 1a, top rectangle wider in Fig. 1b, right circle larger in Fig. 1c) were counted as +1, judgments in the opposite direction (e.g. right T-line lower in Fig. 1a) as −1, and judgments of 'equal' as zero. Each O's score was the algebraic sum of his 12 separate T-judgments and could thus vary from +12 to −12.

The individual scores for the experimental Os are presented in Table 1. The magnitude of the after-effect increased with length of I-exposure and decreased with length of T-exposure. An analysis of variance showed both effects to be significant at the 5% level, with no interaction between them. Variance attributable to I-T interval did not approach significance.

The scores for the control Os, in each case one-half the sum of O's judgments on the two runs through the T-series, were: +1, −2, +4, −2.5, −2.5, and +1 (in order of increasing T-exposure). With a mean of −0.17 and an SD of 2.62, the control level was not significantly different from zero ($t = 0.159$, $p > 0.80$), which indicates that there was no systematic tendency to distort the T-figures when they were not preceded by the I-figures. The control scores were not included in the analysis of variance because of the smaller Ns.

A comparison was made between the judgments of the control group and the

judgments of the 16 experimental Os who had been exposed to both I- and T-figures for the shorter periods (0.25 or 0.75 sec.). The combined mean for these Os was 2.94, significantly larger than the control mean ($t = 2.45$, $p < 0.025$), indi-

TABLE I

TOTAL AFTER-EFFECT FOR EACH O

T-exposure (sec.)	I-exposure (sec.)					
	3-sec. interval			10-sec. interval		
	0.25	0.75	5.0	0.25	0.75	5.0
0.25	4	7	2	2	6	4
	4	4	8	−1	6	7
0.75	2	6	10	1	3	4
	4	1	2	0	−2	6
5.0	1	1	4	2	2	3
	−1	−3	6	3	0	1

cating that the expected after-effects were obtained even at these short exposures. The mean for the 12 Os with the 5-sec. I-exposures was 4.75, which differed, as expected, much more significantly from the control level ($t = 3.67$, $p < 0.005$).

Due to the somewhat fortuitous selection of the stimuli, no attempt was made to analyze the results in terms of the effects of different stimulus-characteristics. It might be noted, however, that while the tendency toward after-effect varied from figure to figure, this tendency exceeded both the chance expectation and the control group's tendency for 11 of the 12 T-figures.

Is it possible that the obtained after-effects were the result of some kind of guessing habit? A comparison between the mean after-effect obtained using the four pairs of figures involving judgments of both I- and T-figures with the corresponding mean for the eight pairs for which only the T-figure was judged showed no significant difference (the tendency toward after-effect was actually a bit greater in the latter case; $t = 0.52$, $p > 0.60$). This result indicates that Os were not merely alternating their judgments, i.e. judging each T-figure with the category opposite to the one previously used. The development of any such tendency should have been prevented by the random introduction of the five extra figures between some of the successive pairs.

No attempt will be made here to discuss the implications of these results for physiological theories of figural after-effects. On the purely behavioral level, they appear to be consistent with previous results. The principle of increased after-effect with increased I-exposure seems well-established,[7] and this relationship appears to hold for briefer exposures than any thus far reported. Krauskopf's finding that increased T-exposure works in the opposite direction also is supported. It should be noted that Krauskopf's T-exposures (0.3 to 1.5 sec.) were extremely brief relative to his I-exposures (repeated 3-min. presentations). The first several seconds of each exposure thus appear to be of primary importance since after-effects will quickly weaken during the T-exposure regardless of the duration of the I-exposure.

[7] Gibson and Radner, op. cit., 453-467; Hammer, op. cit., 337-354.

The lack of interaction in this experiment seems consistent with this interpretation. The failure to obtain significant differences in after-effect associated with the relatively large differences in I-T interval suggests that this interval may be less important than the regular exposure-periods. If Os had been permitted to fixate upon symmetrical figures (including border lines) during this interval the distorting effects of the I-figure would presumably have been more drastically reduced.

None of these results is inconsistent with Gibson's original description of figural after-effects as examples of perceptual contrast.[8] An extension of Gibson's approach might be made to cover, on a behavioral level, the various after-effects which have subsequently been described. The essential point would be that both the I- and T-exposures affect O's judgments by establishing and modifying norms for stimulation in different parts of his visual field. When more is known about how different characteristics of the stimulus-objects influence the judgments (*e.g.* the operation of the distance paradox[9]), a quantitative description of figural after-effects should be possible. This description might be expected to follow the general lines of Helson's or Johnson's approach, in which each stimulus, past or present, has a weighted effect upon O's judgments.[10] The results of the present experiment suggest that the weighting would have to be sensitive to very small differences in the duration of exposure.

Summary. Visual figural after-effects were obtained with I- and T-exposures of less than 1 sec. The magnitude of after-effect increased with the duration of the I-exposure and decreased with the duration of the T-exposure. Variation of the I-T interval had no influence on the magnitude of the after-effects.

---

[8] Gibson, Adaptation, after-effect and contrast in the perception of curved lines, *J. exp. Psychol.*, 16, 1933, 1-31; Gibson, Adaptation, after-effect, and contrast in the perception of tilted lines: II. Simultaneous contrast and the areal restriction of the after-effect, *ibid.*, 20, 1937, 553-569.

[9] Köhler and Wallach, *op. cit.*, 269-357; for difficulties here, see B. H. Fox, Figural after-effects: Satiation and adaptation, *J. exp. Psychol.*, 42, 1951, 317-325.

[10] Harry Helson, Adaptation-level as a basis for a quantitative theory of frames of reference, *Psychol. Rev.*, 55, 1948, 297-313; D. M. Johnson, *The Psychology of Thought and Judgment*, 1955, 324-367.

EFFECT OF THE RELATIVE VOLUME OF STANDARD AND COMPARISON-OBJECT ON HALF-HEAVINESS JUDGMENTS

By Richard M. Warren and Roslyn P. Warren, Brown University

In two recent experiments on weight-scaling, comparison-weights were of the same shape and size as the standards, and weights chosen as half as heavy were more than half the physical weight of the standard.[1] If the comparison-weights were presented in containers smaller than the standard, it would be expected (in view of the size-weight illusion) that their apparent weight would be relatively greater, and the positive deviation of the psychological scale from the physical scale would be less. Individuals probably are familiar with conditions under which an object known to be half the weight of another occupies half the volume (*e.g.* when the weight of one member of a pair of similar objects is compared with both together).

TABLE I

WEIGHTS AND DIMENSIONS OF STANDARDS AND COMPARISON-OBJECTS

Standard (gm.)	Comparison (gm.)	Dimensions (cm.)	
		Type 1/2	Standard & Type 1
20	8, 10, 12, 14, 16, 18	2.6×4.6×1.1	2.6×4.6×2.2
40*	8, 12, 16, 20, 24, 28	2.6×4.6×2.2	2.6×4.6×4.4
80	32, 40, 48, 56, 64, 72	2.6×4.6×4.4	5.2×4.6×4.4
140	28, 42, 56, 70, 84, 98	4.5×7.0×2.6	4.5×7.0×5.2
200	100, 115, 130, 145, 160, 175	4.5×7.0×5.2	9.0×7.0×5.2
320*	96, 128, 160, 192, 224, 256	4.5×7.0×5.2	9.0×7.0×5.2

* Type 1/4 was used with these standards. The dimensions of the Type 1/4 comparison-objects were, for 40-gm. standard, 2.6×4.6×1.1 cm., and, for the 320-gm. standard, 4.5×7.0×2.6 cm.

It was considered that if experience with stimulus-dimensions provided the basis for half-heaviness judgments, the psychological and physical weight-scales should coincide under familiar stimulus-conditions—that is, when the object having half physical weight occupied half the volume of the standard. The present study was designed to test this hypothesis.

Method. Six weighted pill-boxes served as standards. For each standard, an appropriate series of six comparison-weights was constructed in boxes identical in volume with the standard (Table I). These series will be called *Type 1*. Duplicating the weights of these comparison-objects was another set, called *Type ½*, with

volumes half that of the standards. For two of the standards (40 and 320 gm.) there was a third comparison-set, *Type ¼*, with volumes one-quarter that of the standards. The stimuli were made from two sizes of glazed, white, rectangular, cardboard pill-boxes: 2.6 × 4.6 × 1.1 cm. and 4.5 × 7.0 × 2.6 cm. All stimuli consisted of one or more boxes of the same size held together with adhesive tape. The boxes, weighted with copper one-cent pieces and iron paper-clips, were packed with cotton to prevent rattling.

Twenty-four *O*s, 9 men and 15 women undergraduate students, were employed with Types 1 and ½. Each *O* was presented with a card which had the following typewritten instructions:

You are to choose from the comparison-boxes one which feels half as heavy as the standard. You may use either your right or left hand for lifting the boxes but you must use the same hand throughout the test. Lift only one box at a time. You may lift them as many times as you wish.

O completed experiments of Type 1 and Type ½ in one session, which was usually about 45 min. long. *O* was seated at a table in a dimly lit booth facing the seated *E*. An opening in the front of the booth permitted *E* to observe *O*'s manipulation of the boxes, but *O* could not see whether *E* was watching or not.

Some *O*s lifted and held the boxes with their palms above and their fingers grasping the sides. Others allowed the boxes to rest on top of their upturned palms. These different techniques did not seen to cause any difference in judgment. Four different methods of presenting Types 1 and ½ were used. Half the *O*s started with Type 1 and half with Type ½. Half the *O*s starting with each type were presented with standards in order of ascending weight, while the other half received standards in descending order.

After 16 of the *O*s had completed their judgments with Types 1 and ½, they were presented with Type ¼. Each of the four methods of presenting Types 1 and ½ was represented by 4 *O*s. The 8 *O*s who had received standards in order of ascending weight received the 40-gm. standard with Type ¼ before the 320-gm. standard. The other 8 *O*s received the 320-gm. standard first.

Results. The results for Type 1 and Type ½ are summarized in Table II. The difference in estimates of half-heaviness between these two types was significant beyond the 1% level for each of the standards. The results obtained for Type ¼ are presented in Table III. The values determined with the same standards for Types 1 and ½ are included for the purpose of comparison.

It is apparent that the size of the comparison-weight affects the half-heaviness judgment. The larger the comparisons relative to the standard, the greater the physical weight judged half as heavy as the standard.[2] Thus, when the comparisons were the same size as the standard (Type 1), weights more than half the physical weight were chosen. When the comparisons were half the volume of the standards (Type ½), comparisons having about half the physical weight were judged half

---

[2] Only one *O* failed to show this relationship. His judgments for both Types 1 and 1/2 were approximately 50% of the standard. This student had experience in estimating the different weights of boxes having the same volume. He worked for a firm that shipped books, and he was required to estimate weight in connection with the determination of freight-charges. His data are not included.

as heavy. When the size of the comparisons was reduced further to one quarter that of the standard (Type ¼), the half-heaviness selections were less than half the physical weight of the standard.

The role of the relative size of standard and comparison-objects in such experiments has been eliminated by Joy who had the Os lift hidden containers.[3] The weight of the comparisons could be varied continuously by adding or removing water. Thus, the judgments of the Os were affected neither by a predetermined

TABLE II

HALF-HEAVINESS JUDGMENTS (GM.) OBTAINED WITH TYPE-I AND TYPE-I/2 COMPARISONS

(Percentage-values represent the fraction of the standard weight chosen as half as heavy.)

Standard (gm.)	Geometric mean		Median	
	Type 1/2	Type 1	Type 1/2	Type 1
20	9.998 (49.99%)	13.13 (65.65%)	9.909 (49.54%)	13.00 (65.00%)
40	20.84 (52.10%)	24.40 (61.00%)	20.80 (52.00%)	25.00 (62.50%)
80	43.33 (54.16%)	53.51 (66.89%)	44.73 (55.91%)	55.08 (68.85%)
140	68.70 (49.07%)	86.81 (62.01%)	70.87 (50.62%)	91.00 (65.00%)
200	113.5 (56.75%)	134.5 (67.25%)	111.8 (55.90%)	139.0 (69.50%)
320	170.7 (53.34%)	208.8 (65.25%)	173.1 (54.09%)	213.8 (66.81%)
Combined	(52.5%)	(64.7%)	(53.0%)	(65.9%)

TABLE III

HALF-HEAVINESS JUDGMENTS (GEOMETRIC MEAN IN GM.) OF STANDARDS EMPLOYED WITH ALL THREE COMPARISON-GROUPS

(Percentage-values represent the fraction of the standard weight chosen as half as heavy.)

Standard (gm.)	Type 1/4	Type 1/2	Type 1
40	16.08 (40.20%)	20.84 (52.10%)	24.40 (61.00%)
320	134.2 (41.94%)	170.7 (53.34%)	208.8 (65.25%)

range of weights nor by discrete steps within the range as in previous experiments. Joy's results showed judgments of half-heaviness clustering about the physical half-weight (median value 50.9%). The judgments obtained with Type ½ in the present experiment also cluster about the physical half-weight (median value 53.0%). It may be concluded, therefore, that scales of subjective weight based on half-heaviness judgments approximate the physical weight-scale when (1) comparison-objects are one-half the volume of the standards, or (2) cues relating to the size of standard and comparison-objects are absent. The results of the various experiments are summarized in Table IV.

Table V shows the deviations in density of the half-heaviness judgments obtained with the three experimental groups of the present experiment. With Types 1 and ¼, comparison-objects selected by the Os are closer to the density of the standard than the comparison-object having half the physical weight. In Type ½, the standard and the object weighing half as much have the same density, and here we find selection of comparison-objects close to physical half-weight.

If we consider that Os anticipate that an object having half the weight of another

---

[3] Personal communication from F. J. Dudek and K. E. Baker.

should also occupy half the volume, then we may regard the deviation of the *VEG* scale from the physical scale on the basis of set. Thus, any appreciable difference in

TABLE IV

COMPARISON OF CONDITIONS AND RESULTS IN WEIGHT-FRACTIONATION EXPERIMENTS

	Warren & Warren		Harper & Stevens	Guilford & Dingman	Joy
	Type 1/2	Type 1			
Relative volumes of standard (S) and comparison-objects (C)	C = 1/2S	C = S	C = S	C = S	C & S not visible
Range of S-densities	0.61–0.98	0.61–0.98	0.30–3.8	0.52–9.0	C & S not visible
Range of C-weights for each S (heaviest C/lightest C)	1.7 –3.5	1.7 –3.5	1.4 –2.5	10.0	C continuously variable
Range of percentages of S judged half (medians)	49.5–55.9	62.5–69.5	64.5–79.0	51.6–60.3*	48.8–54.3
Median value of range	53.0%	65.9%	68.7%	57.5%	50.9%

* These values are derived from geometric medians calculated by finding the antilogs of medians of log-weight of comparisons judged half as heavy as the standard.

TABLE V

RELATIVE DENSITY (%) AND HALF-HEAVINESS JUDGMENTS

(Density of the standard is taken as 100%)

	Comparison-object	
Type	Half physical weight	Judged half as heavy
1	50	65
1/2	100	105
1/4	200	164

the densities of standard and comparison-stimuli would be contrary to the expectations or set of the Os. The judgments actually made decrease and minimize these density-differences.

Summary. The effect of the relative sizes of standard and comparison-weights upon half-heaviness judgments was studied. Three groups of comparison-objects, which differed in volume but not in weight, were employed. When the comparison-objects were half the volume of their standard, the weights judged half as heavy were about half the physical weight of the standards. When the comparison-objects were the same volume as the standards, the weights judged half as heavy were significantly greater than half the physical weight of the standards. When the comparison-objects were one-quarter the volume of the standards, the weights judged half as heavy were significantly less than half the physical weight of the standards. The results indicate that the relative densities of objects systematically influence estimates of half-heaviness. It is suggested that unfamiliar density-relationships are involved in the deviation of the *VEG*-scale from the physical weight-scale.

THE EFFECT OF PARTIAL REINFORCEMENT AND LENGTH OF ACQUISITION-SERIES UPON RESISTANCE TO EXTINCTION OF A MOTOR AND A VERBAL RESPONSE

By Donald J. Lewis and Carl P. Duncan, Northwestern University

The effects of partial reinforcement are known as the most powerful and ubiquitous.[1] The present experiment was an attempt to study two different responses in a single experiment, a verbal response expressing an expectancy and a motor response continuing the experiment, and to vary reinforcement and length of acquisition-series simultaneously. A previous study indicated that an 8-trial acquisition-series was adequate to demonstrate partial reinforcement effects.[2] Therefore, acquisition-series of 4, 8, and 16 trials were selected for this study.

METHOD

Subjects and apparatus. The Ss were 180 volunteers drawn from introductory psychology classes at Northwestern University. The major piece of apparatus was an ordinary deck of playing cards.

Experimental design. The Ss were assigned to the various groups in sequence as they reported. Two percentages of reinforcement—50% and 100%—were manipulated factorially with three lengths of the acquisition-series—4, 8, and 16 trials.

Procedure. The Ss were told to play a kind of poker game in which each hand was made up of four cards. The Ss themselves dealt the four cards face up and then guessed whether the hand was a 'winner' or not. If the hand was a winner, they were immediately given a nickel. If it was a loser, they were given nothing. They played through the deck, shuffled, and began again. They were told to play as long as they wanted, and that when they were through, they could keep all the money they had won. They were also told that their guesses had nothing to do with the payoffs, which were to be determined solely on the basis of the hand values. Actually, payoffs during acquisition were determined by a prearranged schedule according to the experimental group to which S was assigned, and, of course, there were no payoffs during extinction. The total number of plays for each S (motor response) and his expectation of win or lose (verbal response) were recorded.

RESULTS AND DISCUSSION

Motor response: Continuation of play. The number of 'extinction' plays for each S was determined and converted to common logarithms. An anal-

ysis of variance was performed on the transformed data. Table I gives
the mean logarithms of plays-to-extinction for the various groups. The F
associated with percentage reinforcement was 1.32 (not significant), al-
though the trend of the data was in the expected direction, and for the
length of the acquisition series it was 4.23 (p at approximately 5% level

TABLE I

SUMMARY OF EXTINCTION DATA FOR BOTH VERBAL AND MOTOR RESPONSES

	Mean log number of responses to extinction	F	Mean arc sin % of positive guesses during extinction	F
50%	1.62	1.32	29.47	5.34*
100%	1.55		34.19	
4 Trials	1.68		30.18	
8 Trials	1.61	4.23*	3.125	1.29
16 Trials	1.47		34.06	
%×Trials		1.55		—

* Significant at 5% level.

for 2 and 174 df.). Inspection of the data shows that the group with 4 wins
played the longest, the 8 reinforcement group the next longest, and the
16 reinforcement group quit the most quickly.

Verbal response: Statement of expectancy. The percentage of 'win' expectancies
was determined for each S and an analysis of variance performed on an arc sin
transformation of the percentage data. The right hand columns of Table I show the
means and Fs for the various groups. The effect of percentage reward on the per-
centage of win-guesses was significant at the 5% level for 1 and 74 df., with the
100% reward group giving more positive expectancies during extinction than the
50% group. The effect of the number of acquisition trials was not significant.

The mean percentage of positive guesses was vincentized in sixths through the
extinction-series and in fourths through the acquisition-series. The resulting curves
for the percentage reinforcement variable are shown in Fig. 1. It can be seen that per-
centage reinforcement had a decided effect during acquisition, with 100% reward
resulting in a greater number of positive guesses. It is also clear that an asymptote
of acquisition had not been reached when extinction started.

The typical finding concerning the effects of partial reinforcement upon guessing
behavior is that 50% reward results in a greater number of positive responses in
extinction than does 100% reward.[3] Our findings show exactly the opposite and
have value in suggesting those conditions under which partial reinforcement is not
effective. There are perhaps three major reasons for this reversal: (a) the acquisition-
series was too short for the differential effects of 100% and 50% reward to take
place. The Ss of Grant, Hornseth, and Hake for example, were given 30 additional

[3] Jenkins, *op. cit.*, 222.

trials *after* an asymptote was reached;[4] (b) the stimulus-materials presented to *S* had too many built-in associations. Some *S*s became absorbed in proving or disproving their own hypotheses about winning patterns of cards and seemed to pay little attention to the money reinforcements; (c) the length of the extinction-series was perhaps too short for the 100% group to drop below the 50% group. Because
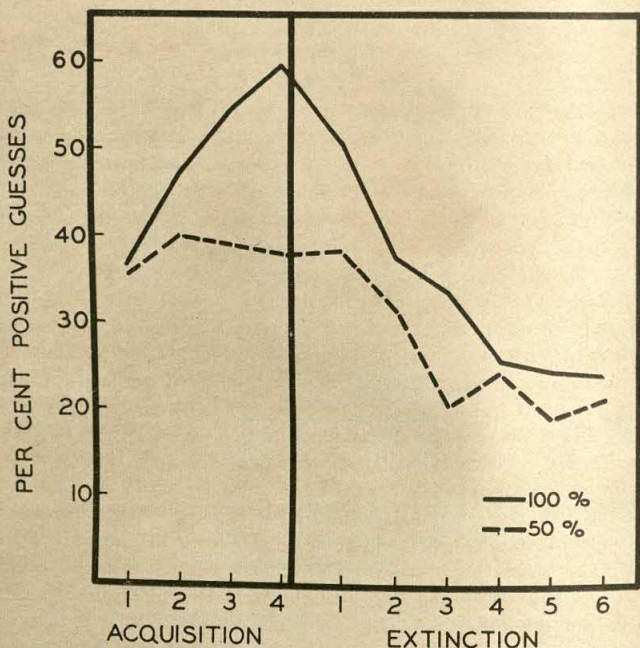
FIG. 1. CURVES SHOWING THE EFFECT OF PERCENTAGE REINFORCEMENT UPON PERCENTAGE POSITIVE GUESSES OF WIN IN ACQUISITION AND EXTINCTION

the length of the extinction-series was self-paced, as compared to the usual guessing study where *E* determines the length of the extinction-series, many *S*s were able to quit quite quickly.

One finding of this experiment that may warrant further investigation is that of the effects of the number of acquisition-trials on number of plays to self-paced extinction. Our data show that the longer the acquisition-series, the fewer the number of extinction-plays. This may be interpreted as due to stimulus-satiation or, its motivational equivalent, boredom.

[4] D. A. Grant, J. P. Hornseth, and H. W. Hake, The influence of the intertrial interval on the Humphreys' random reinforcement effect during the extinction of a verbal response, *J. exp. Psychol.*, 40, 1950, 609-612; D. A. Grant, L. M. Schipper, and B. M. Ross, Effect of intertrial interval during acquisition on extinction of the conditioned eyelid response following partial reinforcement, *J. exp. Psychol.*, 44, 1952, 203-210.

A FURTHER INVESTIGATION OF VARIOUS STIMULUS-OBJECTS IN DISCRIMINATIVE LEARNING BY CHILDREN

By Allen D. Calvin, J. J. Clancy, and J. B. Fuller,
Michigan State College

In a recent paper Calvin and Clifford reported that it was more difficult for children to discriminate between different hues than to discriminate between achromatic stimuli, different brightness-levels of the same hue, or pattern-stimuli.[1] The present experiment is an additional study on the same subject, utilizing other stimulus-pairs.

Procedure. Inasmuch as the present experiment closely follows the previous experiment, only those details of procedure are given here that differ from the experiment of Calvin and Clifford.

Subjects. The Ss (40 pupils from a local public school) were divided randomly into five groups of eight each, except for the restriction that each group have four Ss from the first grade and four from the second. Group I had to discriminate between a blue equilateral triangle with 2-in. sides and a red equilateral triangle with 2-in. sides.[2] Group II had to discriminate between a black equilateral triangle with 2-in. sides and a white equilateral triangle with 2-in. sides. Group III had to discriminate between two black equilateral triangles with 2-in. sides, one of which had the point facing up and the other the point down. Group IV had to discriminate between a black equilateral triangle with 2-in. sides and a black equilateral triangle with 3-in. sides. Group V had to discriminate between a black equilateral triangle with 2-in. sides and a black square with similar area. A balanced experimental design was used with all groups, *i.e.* half the Ss in Group I had red positive and the other half blue positive, etc.

At the conclusion of the testing session each S was asked the following questions: (a) How did you know which cup the toy was under? (b) Did you notice anything besides the two cups? [If the S answered (b) affirmatively, question (c) was asked.] (c) What did you see?

Results. Table I presents median trials to learn, and the percentage of Ss in each group who solved the problem. As can be seen from the table, the performance of Group I, the red vs. blue group, is the poorest. Utilizing the Mann Whitney U-test, we find that the difference between Group I and Group IV is significant beyond the 1% level, and the difference between Group I and the other groups combined is significant beyond the 5% level. This finding is in agreement with the results obtained by Calvin and Clifford, who found color least readily learned by children.[3] It is in contra-

[1] A. D. Calvin and L. T. Clifford, The relative efficacy of various types of stimulus-objects in discriminative learning by children, this JOURNAL, 69, 1956, 103-106.
[2] Our red approximates Oriental Red No. 130 and our blue approximates Italian Blue No. 82 according to the color chart presented in the unabridged second edition of Webster's *New International Dictionary*, 1949, 528.
[3] Calvin and Clifford, *op. cit.*, 105.

diction to the findings of other investigators who used sub-human primates, and found that color differences were the *easiest* to distinguish.

Gellermann[4] and Weinstein,[5] who compared the performance of sub-human primates and children, both suggested that the children were utilizing symbolic processes which enabled them to form generalizations that were of considerable aid in discriminative learning. This is in line with the explanation of Calvin and Clifford, who proposed that the general concept *colored cards* served in this case as a handicap by preventing the Ss from using the different hues as cues.

On the basis of the Ss' responses to the questions which were asked at the termination of the experimental session, the possibility of an explanation in terms

TABLE I

MEDIAN TRIALS TO LEARN

	Group				
	I	II	III	IV	V
Trials	40	40	20	13.5	28.5
% Ss who solved	25	37.5	62.5	87.5	50

of the difference in hue being subliminal can be ruled out. Our red and blue hues were easily distinguished by the Ss. We therefore feel justified in postulating that the difference in behavior between children and sub-human primates on color discriminative learning problems is primarily due to the fact that the conceptualization process which acts as an aid in many cases this time acted to handicap the children.

The findings presented in Table I also suggest that the size-discrimination of Group IV was the easiest problem. Group IV's performance is not only better than that of Group I, but it is also significantly better ($p < 0.05$) than that of Group II which had to solve a black vs. white problem.

As Huang[6] has demonstrated in a matching experiment with children, and Nissen and Jenkins[7] suggest in their work with chimpanzees, it is undoubtedly possible to vary the results obtained by making the difference in size extremely pronounced relative to the difference in hue, or conversely making the difference in hue extremely pronounced relative to the difference in size. We doubt, however, that this will serve as an adequate explanation for our results. The ratio of large area to small area which we utilized closely approximates that used by experimenters who worked with sub-human primates. The relatively small size-difference seems phenomenologically much less striking than the marked difference between hues used as the cues for Group I or the black and white cues used by Group II. If this factor is operating, it should tend, therefore, to favor Groups I and II and would work toward cutting down the relative inferiority in performance of these groups when compared with

[4] L. W. Gellermann, Form discrimination in chimpanzees and two-year-old children: I. Form (triangularity) per se, *J. genet. Psychol.*, 42, 1933, 3-27.

[5] Benjamin Weinstein, Matching-from-sample by rhesus monkeys and by children, *J. comp Psychol.*, 31, 1941, 195-213.

[6] I. Huang, Abstraction of form and color in children as a function of the stimulus-objects, *J. genet. Psychol.*, 66, 1945, 59-62.

[7] H. W. Nissen and W. O. Jenkins, Reduction and rivalry of cues in the discriminative behavior in chimpanzees, *J. comp. Psychol.*, 35, 1943, 85-95.

Group IV. We feel that this is a real phylogenetic difference which again is ascribable to the children's conceptual ability with the concepts 'large' and 'small' in this case acting as aids to the solution of the problem as opposed to the handicap created by the concept 'colored' which was discussed in the previous study.

Summary

The relative difficulty of various stimulus-dimensions for discriminative learning by children was studied. It was found that it was most difficult to discriminate between stimuli varying in hue and least difficult to discriminate between stimuli varying in size. The relationship of these findings to the results of experimenters who worked with sub-human primates is discussed.

RECOGNITION AS A FUNCTION OF MEANINGFULNESS
AND INTENTION TO LEARN

By Robert L. Karen, San Diego Junior College

In a recent experiment by Postman, Adams, and Phillips, the recall of intentional learners exceeded that of incidental learners for materials of low but not for high associative value.[1] The recall of both groups increased with associative value. When learning was tested by the method of recognition, however, no such relationships appeared. The experiment to be reported was designed to provide further data on recognition as a function of meaningfulness of material and intention to learn.

Subjects. Thirty psychologically naïve volunteers from adult-education classes, ranging in age from 17 to 50 yr., were divided into two equal groups, one of intentional and the other of incidental learners. The Ss were assigned at random to the two groups, except that care was taken to see that there was an equal number of women in each, since women are superior to men in tasks of the type used in this experiment.[2]

Materials. Fifteen nonsense syllables were taken at random from Hull's 1933 list.[3] The syllables selected had associative values ranging from 20-30%. Fifteen three-letter words were selected in the same manner from the Thorndike-Lorge General Word List.[4] The words selected occurred at least once in a million times in reading material. These two sets of stimulus-materials made up the learning list. A practice-list was prepared in the same manner, consisting of five syllables and five words. Each word or syllable in each list was typed on a separate card and in the same position on the card. Each list was then randomized, and the random order thus established remained the same during the experiment.

The orienting task was to find words or syllables in a matrix of all of the items typed in two columns on a sheet of paper—the syllables in one column and the words in the other. The items in each column were in scrambled orders. A similar sheet was prepared with the practice-words and syllables.

For the test of recognition a list was prepared of the 30 original words and syllables plus an equal number of dummy words and syllables. Each of the new items

had the same number of letters and the same first letter as a corresponding item in the original list. The order of the 60 items was randomized and the list was typed in 4 columns of 15 items each. The items were numbered consecutively from 1-60.

Procedure. S was seated comfortably at a table and given the following instructions:

This is an experiment dealing with the ease of locating code words and syllables. I will present to you a series of cards containing code words and syllables. You are to look at the word or syllable, locate it on this list in front of you [E now displays the practice-matrix], and draw a line through it with this pencil. You will be given 15 sec. to look at the card, find the word or syllable on the list, and draw a line through it with your pencil. I will present the cards at a constant rate, so be ready for the next card. I will announce each card by saying 'next.' Let us practice with these cards first.

The practice-stimuli were presented at the rate of one every 15 sec. (found in a pilot-study to be satisfactory even for the slower Ss). When the practice-task was completed, E asked S if there were any questions about the task, and answered any questions asked. Then, before the first trial, the intentional Ss were instructed to "study and remember each word or syllable" presented with view to recognizing them in "a test later on in the experiment." The cards were then presented, one at a time, with each card placed upon the other, until all 30 had been exposed. After the last card had been presented and all materials concealed, the time was noted and E announced a 1-min. rest period, following which he read these instructions:

You have been presented words and syllables on these cards. Now I will give you a sheet of paper containing a list of 60 words and syllables. Thirty of these you have already seen on the cards, 15 of which were the code-words and 15 of which the code-syllables. The remaining 30 are decoys. Your job is to look over the list and draw a line through the words and syllables you recognize as being on the series of cards I have just shown you. Remember, there are 15 correct words and 15 correct syllables."

This passage was practiced aloud by E until he was able to read it at a slow but constant rate, giving each S the same delay (40 sec.) before the test of recognition. Five minutes were allowed for the test. Raw scores were corrected for guessing by the use of the following formula: Corrected Raw Score = Number Right $-$ $\frac{1}{2}$ Total Number of Responses. Each S was given a corrected score for the words and a corrected score for the syllables.

Results. The results of the experiment are presented in Table I. The performance of the intentional learners was superior to that of the incidental learners although the superiority was based largely on recognition of words. The difference between the two groups in terms of syllable-recognition was not statistically significant. Both groups recognized more words than syllables, but the word-syllable difference was significantly greater for the intentional group than for the incidental group. In the experiment of Postman, Adams, and Phillips, no significant differences in recognition as a function of intent or materials were found.[5] Nor is it possible merely on the assumption of the greater sensitivity of recognition to account for the discrepancy between the earlier results for recall and the present results for recognition. The

---

[5] *Op. cit.,* 1-10.

intentional learners of Postman, Adams, and Phillips surpassed their incidental learners in the recall of items of low associative value, but not in the recall of items

TABLE I

RECOGNITION OF WORDS AND NONSENSE SYLLABLES AS A FUNCTION OF INTENT

	Intentional	Incidental	Diff.	t
Words+syllables	9.47	5.70	3.77	3.89*
Words	5.67	3.13	2.54	4.60*
Syllables	3.80	2.57	1.23	1.68
Words−syllables	1.87	0.56	1.31	2.33†
t	4.64*	2.38†		

* Significant beyond the 1% level of confidence.
† Significant beyond the 5% level of confidence.

of high associative value, while precisely the opposite results appeared here in recognition. Procedural differences in the two experiments were, of course, numerous.

Summary. An experiment is reported on recognition of verbal materials as a function of meaningfulness and intention to learn. The results are contrasted with those of an earlier experiment by Postman, Adams, and Phillips.

CONTEXT IN THE PERCEPTION OF SENTENCES

By ELMO E. MILLER, University of Minnesota

This investigation of inter-sentence context extends the work of earlier investigators on smaller units of context. Cattell found that more letters could be perceived if organized in meaningful words,[1] and Pillsbury found that responses to words in a tachistoscope depended, in part, upon preceding words in the series.[2] The present experiment demonstrates that more of a sentence can be perceived when successive sentences tell a story than when the sentences are in random order.

Method. Each of 50 sentences was presented twice to each of 16 Os, once in an order constituting a simple story, and once in random order. The sentences were exposed for 0.15 sec. in a Dodge mirror-tachistoscope, and the Os wrote immediately what they could of each sentence. The presentation of the sentences was prefaced by standardized spoken directions. The only hint that the conditions might vary within the experiment was given by the statement. "In some of them, you will notice that the sentences seem to tell a story; in others, you will not notice such a relation." The Os were not told which parts of the series of sentences would be in a storied relation.

The Os were given 10 practice trials on other material to adapt them to the apparatus and to the experimental conditions. The sentences used for the experiment proper were on a simple level, typical of those given children who are learning to read. This level was chosen to insure highly overlearned perceptual habits, and to increase redundancy to the point at which the Os would get a large part of the meaning of a sentence even if they managed only to report half of it. Half of the Os were presented first with 25 sentences in meaningful order, following which there was a short rest-period. Then they were given 25 sentences in random order, a slightly longer rest-period, 25 more sentences in random order, another short rest-period, and the remaining 25 sentences in meaningful order. For the other half of the Os, the presentation was reversed, with the 50 sentences appearing in meaningful order in the middle of the sequence and in random order at beginning and end.

There were two sets of typewritten stimulus-cards, as nearly identical as possible, and each set was used half the time for each condition. The preëxposure fixation-point was placed at a point corresponding to the middle of the typical sentence.

Results. For 3 of the 16 Os, both more words and more letters were perceived when the sentences were presented in random order; in a fourth case, more letters but fewer words were perceived; in a fifth, fewer letters but more words were perceived. (Credit was given for all words recorded regardless of order, and for

letters contained within those words.) By Wilcoxon's nonparametric test for paired replicates,[2] the difference in favor of the meaningful order was significant, for letters, beyond the 5% level, and for words, beyond the 2% level.

An index which weighted both letters and words about equally also was computed —the product of the number of words recorded and the average number of letters per word in the material used was added to the number of letters recorded. This index, which resolved conflicts for the two cases in which differences for letters and words were in opposite directions, showed the superiority of the meaningful order at a level of confidence beyond 2% (Wilcoxon's test). The three Os whose combined letter-word scores were not greater for the meaningful order tended to be high scorers, ranking first, fourth, and seventh in the group of 16 Os. Their performance was so proficient to begin with, apparently, that the meaningful order of presentation could make little contribution to their scores. For Os performing near the 'ceiling' of the test, differences due to order probably represented random fluctuations.

Conclusion. The tachistoscopic perception of sentences is facilitated when the sentences are presented in a meaningful as compared with a random sequence.

---

[2] Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures,* American Cyanamid Company, 1949, 1-16.

APPARATUS

AUTOMATIC MEASUREMENT OF GENERAL ACTIVITY IN TIME-UNITS

By Edward Newbury, University of Kentucky

Two reviewers of the literature on spontaneous activity have remarked on the quantitative inadequacies of records obtainable from 'stationary' activity cages.[1] With various types of apparatus sensitive to general movements of either human or animal Ss, however, it has been possible to obtain from the movement-records useful measures in units of time.[2] One of the methods of quantification has been to segment an experimental record into equal time-intervals and then to determine the number of these in which any activity occurs. This method at best is tediously impractical, or, as in the hands of Hunt and Schlosberg who used 5-min. intervals, involves units so gross as to distort the description of behavior-patterns.[3]

The writer has developed a simple magnetic-relay system which automatically (1) describes equal time-intervals recurring at periods which may be less than 0.5 sec. (2) indicates on a permanent record which of these intervals contain activity, and (3) supplies on a counter a running total of the number of such 'active' intervals. The system can be used with any

[1] J. D. Reed, Spontaneous activity of animals, *Psychol. Bull.*, 44, 1947, 393-412; Mary Shirley, Spontaneous activity, *Psychol. Bull.*, 26, 1929, 341-365.

[2] E. E. Nicholls, A study of the spontaneous activity of the guinea pig, *J. Comp. Psychol.*, 2, 1922, 303-330; H. M. Johnson, chapter 23 on sleep in W. L. Valentine, *Readings in Experimental Psychology*, 1931, 241-291; J. McV. Hunt and Harold Schlosberg, General activity in the male white rat, *J. comp. Psychol.*, 28, 1939, 23-38; O. C. Irwin, Effect of strong light on the body activity of newborns, *J. comp. Psychol.*, 32, 1941, 233-236. Irwin and Weiss (Differential variations in the activity and crying of the newborn infant under different intensities of light: A comparison of observational with polygraph findings, *Univ. Iowa Stud. Child Welfare*, 9, 1934, 141) used half-minute units to quantify a specific type of activity, non-instrumentally observed.

[3] See Hunt and Schlosberg, *op. cit.*, 30 f., for their difficulties in testing the stability of small cycles and of the general daily pattern of activity. Irwin used a smaller unit, "the number of active seconds per minute" obtained from polygraph records, but did not disclose how they were counted.

type of cage, stabilometer, bed, or situation so designed that activity of the animal will produce electrical contact.

Among numerous relay-systems tested, one of the most satisfactory is illustrated in Fig. 1. Points of the cage-contacts are connected either directly or through some relay-arrangement to the leads (F) shown in the upper right. Another set of contacts, at the lower left, are controlled by a timer,
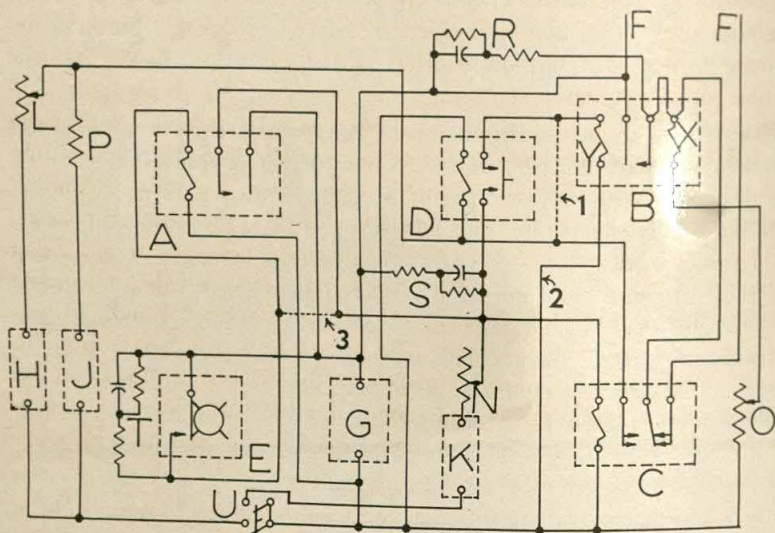
FIG. 1. CIRCUIT DIAGRAM

Specific values and makes of components listed in the key are from a satisfactory experimental version of the apparatus, but are not necessarily fixed or the best. For long runs, where the equipment is unattended, it would probably be safer to use heavier-duty components for relays B and C, and a 12-v. coil in relay C. A = SPST relay, coil 12 v., 100Ω, contacts 8 amp.; B = SPST very fast differential polarized relay, windings (X) and (Y) each 100Ω (Western Electric No. 239 JS, a SPDT relay used as SPST); C = DPDT telephone-type relay, coil 330Ω (Western Electric E-type, double-transfer spring combination, No. D 161966, coil 24 v., a 4-pole relay used as DPDT); D = SPST double-break relay, coil 77Ω, contacts above 8 amp. (Allied Control Co., Circle Relay Div., No. 808030); E = contact-timer operated by 110-v. synchronous clock motor; F = leads to contacts closed by cage-activity; G = DC power-supply, 13.2 v. open circuit, app. 11 v. full load, filtered, rectified AC, rated 6-12 amp. at 12 v. (P. R. Mallory & Co., No. 12 RS 6 D); H = activity signal-marker, 1.5 v., 4Ω; J = electric counter, coil 6 v., 17Ω; K = time signal-marker, 1.5 v., 4Ω; L = variable resistor, 0-200Ω, 4w.; N = variable resistor, 0-75Ω, 4w., in series with fixed resistor, 40Ω, 10 w.; O = variable resistor, 0-60Ω, 4 w.; P = fixed resistor, 10Ω (two 20-Ω, 10-w. resistors in parallel); R = anti-spark assembly: 100,000-Ω, 2-w. resistor in parallel with 2-mfd. condenser (200-v. DC), and in series with a 110-Ω, 2-w. resistor; S = anti-spark assembly: 100,000-Ω, 2-w. resistor in parallel with 2-mfd. condenser (460-v. DC), and in series with 220-Ω, 2-w. resistor; T = anti-spark assembly: 100,000-Ω, 2-w. resistor in parallel with 5-mfd. condenser (300-v. DC), and in series with 220-Ω, 2-w. resistor; U = DPST switch.

at periodic intervals closing and, for a very short duration, remaining closed. With timer-contacts open, those of relay A, controlled by the timer, will also be open, and relay C will be at rest, the connection through its resting contacts being unbroken. In this situation, closing the cage-controlled contacts by the animal activates coil X of relay B, the contacts of which are thereby closed and electrically self-locked. When, at the end of the periodic interval, the timer-contacts close, the contacts of relay A then close and activate relay C, the resting contacts of which are opened, breaking the circuit to the cage-controlled contacts. The apparatus is thereby brought into a period of insensitivity to cage-control without disturbing the locking circuit of coil X in relay B. If the contacts of relay B have already been locked, the apparatus registers, that is, the contacts closed by the activation of relay C activate the counter, the activity-marker, and relay D, the last supplying a time-delay to allow operation of the recording components. Relay B is a fast, sensitive relay of the differential polarized type. Closing contacts of relay D activates coil Y of differential relay B, canceling and reversing the magnetic field locking this relay and thereby opening its contacts and resetting the apparatus. The apparatus is still not sensitive to cage control, however, until, at the close of the timer-controlled refractory period, relay C is released, whereby the lead to the cage-controlled contacts is restored. Periodic time-intervals are recorded by the time-marker hooked in parallel with the coil of relay C.

As can be readily seen, not all components in the diagram are essential, although for practical reasons they may be desirable. If contacts on the timer will carry the current through the contacts of relay A, the latter can be eliminated, provided that the connection is made as shown by the dotted line at 3. Relay C also can be eliminated by building into the timer the double-throw contacts of this relay and an additional pair of contacts to control the time-marker. The remaining two relays operate only when the animal is active, and even one of these (D) can be eliminated if coil Y of relay B can be made to break the circuit of the recording components slowly enough. This was done in tests by inserting a variable resistance in the line at 2 and making the connection shown by the dotted line at 1. With relay D then taken out, the two variable resistances connected with coils X and Y could be so adjusted that the counter and activity-marker operated satisfactorily. In the absence of relay D, the necessary delay is produced by slowing up the speed with which the magnetic field of coil Y is built up sufficiently to open the contacts of relay B. Since this speed changes somewhat with the continued operation of this relay, adjustments of the resistances must be such as to provide sufficient delay when the animal is both minimally and maximally active. The safer procedure is probably to retain relay D. If a differential polarized relay is not available, an additional relay can be substituted for coil Y, which, when activated, will break the circuit through the contacts of relay B.

Experience with various relay-arrangements in developing the present system in-

dicates that for records of long duration incidentally available laboratory components may not be adequate, because cumulative electrical effects in imperfectly selected condensers and contacts can produce inaccurate measurements and an eventual break-down. An advantage of the Fig. 1 design is that the only relays necessary are those which operate when the animal is active. Appropriate selection of components should provide an apparatus satisfactory for studies of long duration. A sustained power-supply, rather than batteries, also has proved necessary.

Functions of the apparatus in relation to units of measurement can be explained, along with some useful terminology, by reference to Fig. 2. The period of experimental observation is divided into intervals we may call cycles (a), defined by the apparatus as the time between the successive openings of the resting contacts of relay C. The cycles are equal to and controlled by the timer. The refractory period (b) is the last part of a cycle
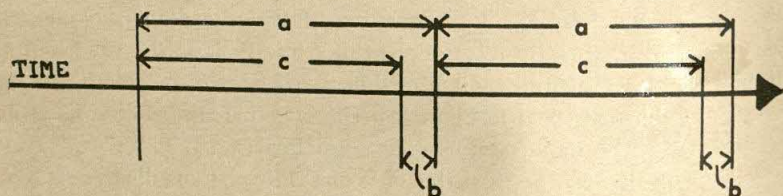
FIG. 2. TIME-RELATIONS
Between cycles (a), refractory periods (b), and 'taks' (c)

and is determined by the length of time the resting contacts of relay C are open. The refractory period is controlled by and closely approximates the duration of closure of the timer-contacts. This period is lost to observation, but can be very brief, 0.1 sec. or less. Preceding the refractory period is that part (c) of the cycle during which the apparatus is sensitive to activity of the animal. This period (the *tak period*) provides the unit of quantification. If the animal has moved sufficiently during this period to activate the apparatus, one *tak* is scored. Thus, with the timer so set that the refractory period is 0.1 sec. and the cycle is 2.0 sec., there are 30 'tak' periods per min., in which a maximum measure of activity of 30 'taks' may be obtained. With a 0.1-sec. refractory period and cycles of 0.5 sec., there would be 120 possible 'taks' per min., each reflecting activity in a period of 0.4-sec. duration.

Since recording takes place during the refractory period, each pulse of the activity marker on the permanent record is separated by at least one 'tak' period. This procedure requires less tape for a distinguishable record than one system tested whereby the bips, recorded immediately on ap-

pearance of the activity, might occur at the end of one and the beginning of the next 'tak' period. In the present system a readily usable 20-min. record at 2-sec. cycles can be obtained on a single kymograph 19 in. in circumference.

The recording system was tested by use of a conventional three-cornered tambour-cage with the standard pneumatic connections to a recording tambour. A section of steel wire was substituted for the writing arm of the latter, and contact was made when, in response to activity of the animal, this wire in its upward swing struck a metallic crossbar. The return torque of such a wire lever was so slight that, probably from arcing, it tended to stick to the crossbar at the contact-point. Consequently these contacts were made to control a standard one-wave thyratron relay operating a fast telephone relay which supplied the cage-controlled contacts referred to in Fig. 1. Since the one-wave thyratron relay is insensitive for alternate periods of approximately 1/120 sec., for research purposes a two-wave thyratron relay would sometimes be preferable, but if enough force is obtainable for control of the cage-contacts, these can be connected directly to the leads from the recording apparatus.

The telephone relay in the thyratron circuit chattered, of course, as it kept pace with the cage-activity. Although the improvement was not incorporated in the apparatus as presently tested, this chattering, which tends to wear out a mechanical relay, can be eliminated by using for the telephone relay in the thyratron circuit a double-pole type which becomes electrically self-locked by the contact of one pole when contact is made with the other pole connected to the recording apparatus. The self-locking pole of the telephone relay must be so connected through an additional pole of relay C of Fig. 1, that the locking circuit is in readiness when relay C is at rest, but is broken when the refractory period is initiated.

AN APPARATUS FOR THE STUDY OF INSTRUMENTAL
LEARNING IN THE RAT

By J. L. McGaugh and Lewis Petrinovich, University of California

The lever-pressing apparatus has been used extensively in studies of instrumental learning.[1] Although the apparatus is useful in studying the variables affecting the performance of a learned response, its suitability for the study of acquisition is questionable. It suffers from two major shortcomings. First, when S is newly introduced into the apparatus, extensive pretraining is necessary to insure learning. This pretraining usually involves several presentations of the reward along with the click of the reward-delivery mechanism. With this procedure, learning occurs so rapidly that the initial phase of learning is not recorded; maximal rate of response usually appears following one or two reinforcements of the response.[2] Furthermore, if, instead of placing S in a strange environment, E introduces the lever into S's home-cage, the initial rate of response prior to reinforcement is so high that there is little room for improvement.[3]

While the lever-apparatus may be more appropriate for the study of performance than for the study of acquisition, it is essential that we be able to record and to control whatever differences in acquisition may occur in the first stages of training. To that end, we have designed an apparatus which permits investigation of the acquisition of a simple instrumental response as well as of its subsequent performance.

The apparatus consists of a rectangular box 24 in. long, 6 in. wide, and 7.5 in. deep (Fig. 1). The floor of the box is separated into two parts by a 4-in. gap. The floor in each end (A and B in Fig. 1) is so hinged that a slight pressure will close the circuit of a microswitch which is underneath. These switches control the feeding mechanism, C, which consists of a table driven by a ratchet-relay.[4] The table is 16.5 in. in diameter and

[1] B. F. Skinner, *The Behavior of Organisms: An Experimental Analysis,* 1938, 48-51.

[2] *Ibid.,* 66-68.

[3] W. N. Schoenfeld, J. J. Antonitis, and P. J. Bersh, Unconditioned response rate of the white rat in a bar-pressing apparatus, *J. comp. physiol. Psychol.,* 43, 1950, 41-48.

[4] The ratchet-relay used in the apparatus is a part of the 'free-game' mechanism of a discarded pin-ball machine and may be obtained from any amusement-machine company.

has 50 small depressions spaced 1 in. apart around the circumference. Any kind of reward, solid or liquid, may be placed in these depressions. The circuit is so designed that the table is automatically turned, thus delivering a reward at D, *only* when A and B are depressed alternately. In other words, when a rat is placed in this apparatus it must learn to shuttle back and forth from A to B to be rewarded. (See Fig. 2 for the wiring diagram.)[5]

The characteristics of the apparatus are illustrated by data for 19 male rats of the $S13$ strain maintained at the University. The 11 Ss of Group I
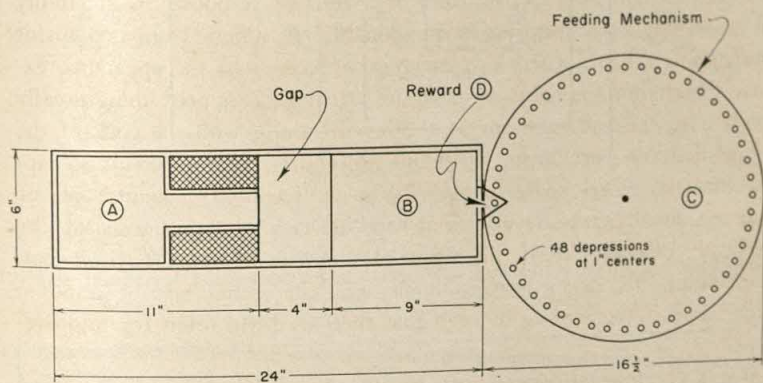
FIG. 1. PLAN OF APPARATUS
Table C, which delivers the reward, is operated by ratchet-relay underneath.
See Fig. 2 for wiring diagram.

were placed on a 24-hr. deprivation-schedule for 4 days. On the fifth day, and on every subsequent day for an additional 13 days, each S was placed in the apparatus for 10-min. during which each shuttle-response from A to B was rewarded. Each reward consisted of about $1/10$ gm. of wet mash. The 8 Ss of Group II (controls) were placed on a 24-hr. deprivation-schedule for 4 days. For the next 7 days, they were placed in the box for 10-min. during which no reward was given. For the following 14 days, each response was rewarded as in Group I. This procedure was used to measure initial response-levels for the apparatus. Each animal was weighed before and after each experimental period and was given enough wet mash to maintain its body weight at about 88% of its satiated weight.

Average daily response-rates for the two groups are presented in Fig. 3. It can be seen that the rate of responding increases from 4.5 to 47.2 for
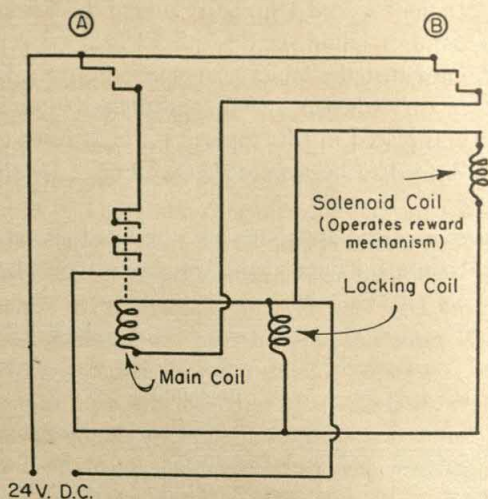
FIG. 2. WIRING DIAGRAM FOR REWARD-DELIVERY MECHANISM

Microswitch B (under floor in reward end of apparatus) operates main coil of relay. Switch A then activates solenoid coil and releases the locked relay. The solenoid is thus operated by closing Switches A and B alternately.
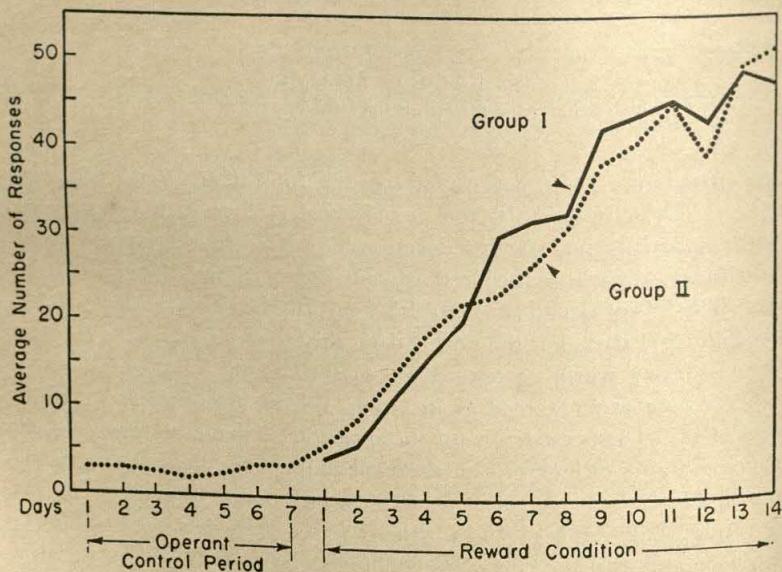
FIG. 3. AVERAGE NUMBER OF RESPONSES DURING DAILY EXPERIMENTAL PERIODS OF 10 MIN. EACH

Group I, and from 5.5 to 51.0 for Group II over the 14-day reward-period. The curve for Group II indicates a low but stable level of response —low enough that the learning process is not obscured but high enough that special pretraining procedures are unnecessary—during the 7-day pre-reward period. Further, there are no differences between the two groups in rate of acquisition,[6] which indicates that initial levels may be determined without affecting subsequent conditioning.

Individual differences in response-rate are extremely stable. The correlation between cumulated responses to Day 3 and Day 14 is 0.88, and, between Day 3 and Day 5, 0.96. The reliability of the apparatus as measured by responses made on odd and even experimental days is 0.99.

The apparatus is versatile enough for work involving partial or periodic reinforcement (the first with only certain of the depressions on the food-tray filled, and the second with an appropriate timing device). It is suitable also for laboratory demonstrations, since an S's performance is not disturbed by the presence of a class, and the occurrence of the response can be signalled easily to the students.

---

[6] The curves of the two groups are representative of the individual curves of the Ss in each group.

A CONTINUOUS DRINKING-RECORDER FOR SMALL ANIMALS

By MICHAEL L. DUFFY and G. E. PRICE, University of Washington

The apparatus was designed to obtain direct, continuous records of drinking over prolonged periods. Hill and Stellar, and Young and Richey have previously developed drinking recorders.[1] The Hill and Stellar apparatus records the animal's contacts with the drinking spout. It was found very useful for recording over short periods of time, but for recording over prolonged periods such as 24 hr., it worked less well. If the recording paper is driven at a speed sufficient to insure accurate readings, the task of translating the records into fluid volume is enormous, and if the recording paper is slowed to a speed that produces a manageable record the error in translating is large. Young and Richey's apparatus utilizes a partial vacuum with an open reservoir of liquid at the bottom to produce the record and it provides an easily read graph of drinking. As they have pointed out, this gives a discontinuous measure of drinking since the level of the fluid in the open reservoir is not constant. We wished to secure continuous records over prolonged periods which could be readily translated with a small error.

Description. The apparatus utilizes a photo-electric cell to follow the fluid meniscus in a stoppered vertical tube. A schematic outline of the mechanical components is shown in Fig. 1. This whole arrangement is attached to a vertical piece of plywood, which is mounted on the face of the cage with the electrical components of the main circuit, shown in Fig. 2, mounted on the back side of the plywood. The drinking tube, made by twirling the end of the glass tube in a hot flame, has an opening of approximately 3/32 in., and the lower 4 in. is bent at a 45° angle. The photo-cell carriage is made of aluminum sheet with the large end, shown on the left in Fig. 1, containing the 930 photo-cell tube and the small end containing the 6-v. light bulb. Current is supplied to the carriage by flexible shielded wire. The drinking tube runs through holes punched in the top and bottom of the carriage. A leaf-spring, 2 in. long, is soldered to a bracket, bolted to the carriage, and presses against the plywood. With a screw adjustment, this provides the pressure of the recording pen on the paper cylinder. The carriage is supported in a continuous loop of heavy fish line which is carried over an idler pulley at the top and the drive-motor
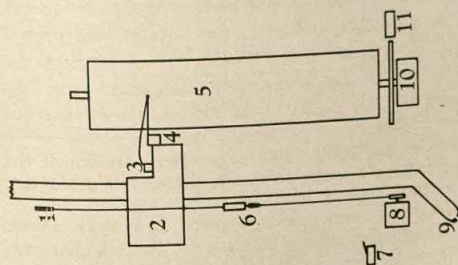
FIG. 1. COMPONENTS OF DRINKING-RECORDER

1—idler pulley; 2—photo-cell carriage unit; 3—ink well; 4—time-marking relay; 5—recording drum; 6—counter balance and tension-spring; 7—limit switch; 8—drive motor and pulley; 9—drinking spout; 10—clock motor; 11—timer circuit microswitch.

MAIN CIRCUIT
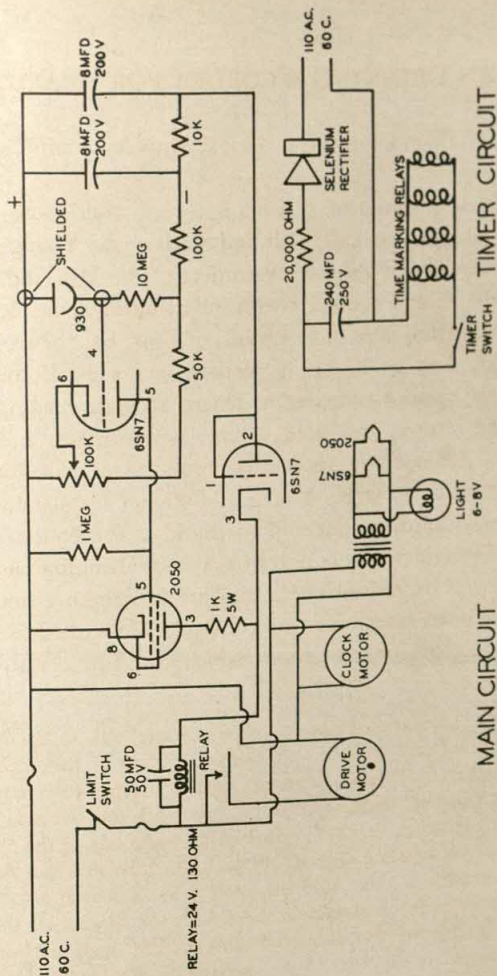
TIMER CIRCUIT

FIG. 2. MAIN CONTROL CIRCUIT AND TIMER CIRCUIT

pulley at the bottom, and is counterbalanced with a weight made from a length of iron bar. The tension spring keeps the drive line tight and yet allows it to be loosened in order to move the carriage manually. The recording cylinders are cardboard tubes driven directly by the clock motor at 1 rph., and an hour-hand drive on one unit turns the notched timer-wheel to operate a micro-switch. The limit switch breaks the main circuit in case the fluid level should drop below the operating range of the recorder.

Operation. The fluid in the tube acts as a focusing lens, and, with proper spacing of the tube, light and photo-cell, the light is focused on the photo-cell when the carriage encounters the fluid-filled part of the tube. The main circuit, shown in Fig. 2, is so adjusted with the 100 K potentiometer that the drive motor is operating when the tube is empty and stops when the fluid focuses the light on the photo-cell.[2] As the animal drinks, the carriage follows the meniscus of the fluid down the tube. The pen arm is soldered to the timer relay and makes a small vertical mark every 55 min. This is necessary for recording over long periods, as the animal may not drink for several hours, and the pen will retrace the same line on every revolution of the recording drum.

The timer circuit is also shown in Fig. 2. The values of the resistor and capacitor were chosen to operate the four relay coils serving as markers. The circuit provides insufficient current to operate the relays continuously with the switch closed. The capacitor is charged while the switch is open and furnishes a momentary pulse when the switch is closed. This differentiating circuit gives a clear single vertical mark.

Four of these recorders have been constructed. We used 14 mm. inside diameter Pyrex tubing for the drinking tubes, and 2⅝ in. × 26 in. cardboard tubes for the recording drums. The recording paper was a good grade, high rag-content paper. A moistened rubber stopper was used to seal the top of the drinking tubes. The recording pens consisted of brass capillary tubing fed by a length of plastic capillary tubing from the ink pot. A rule was scribed that the vertical dimension of the record could be read directly in cubic centimeters, and a plastic triangle was scribed to measure the horizontal dimension in 2-min. intervals. Twenty-four hour records can be broken down into desired time-intervals quickly and easily.

Since this arrangement is sensitive to light it was found necessary to operate the recorders under constant illumination, but this might be avoided by proper shielding of the photo-cell unit. The basic design has considerable versatility. By using drinking tubes of a smaller diameter finer measurement is possible. The units described above responded accurately to changes as small as ⅛ cc. The drinking spout was found to be effective hence there was little or no loss of fluid during drinking. One source of difficulty is the expansion of air in the drinking tube above the fluid surface due to the heat generated by the light-source. This problem was solved adequately by filling the tubes with fluid of room temperature and letting the unit stabilize for at least 15 min. before inserting the drinking spout into the cage. Large fluctuations in line-voltage will also influence the circuit, and if such fluctuations are extreme a constant voltage transformer may be necessary. If opaque fluids are used, the circuit can be altered that a decrease in light on the photo-cell will start the drive motor.

---

[2] Adapted from Gilbert Smiley, Control circuits for industry, *Electronics*, Jan. 1941, 29.

APPARATUS NOTE

MAZES FOR SMALL AQUATIC ANIMALS

The construction of a small, inexpensive, water-tight maze with uniform channels has offered a problem to investigators of the behavior of the protozoans and other small aquatic animals. Mazes etched in glass have gross irregularities in the channels that are detrimental especially for unicellular animals. Small mazes without these irregularities can, however easily be constructed of lucite plastic with only a coping saw, small files, and sandpaper.[1]

The maze-pattern should be cut through a piece of lucite with a width equal to that required for the depth of the maze. This piece is then cemented with plastic
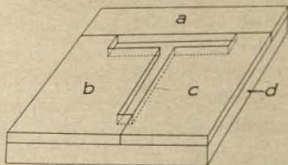
FIG. 1. A SMALL LUCITE MAZE CONSTRUCTED IN SECTIONS

solvent to a lucite base, providing a water-tight maze of uniform depth. If the channels required are no smaller than ⅛ in., the maze pattern may be cut through a single piece of lucite. For smaller mazes, it is advisable, however, to build the maze in sections as greater accuracy may be obtained. Fig. 1 shows a simple maze constructed in sections. Sections *a, b,* and *c* are cemented to a base, *d,* to form this T-maze. To achieve precision in the smaller mazes, the filing, sanding, and cementing should be done under a binocular microscope making careful and frequent comparisons with a measuring instrument. Maze channels as small as 0.5 mm. can be made quite accurately in this manner.

RICHARD H. LAWLESS

University of Wichita

NOTES AND DISCUSSIONS

RECALL OF STIMULUS-ITEMS ARRANGED IN A SQUARE MATRIX

Most studies of the learning of serially presented material are concerned primarily with quantitative measures of the learned material, rather than with the effects of serialization. Such effects have been investigated recently by Anderson and Ross,[1] who studied five different types of stimulus-items (letters, numbers, geometric figures, words, and symbols) arranged in separate, 5 × 5 square matrices. It was found that accuracy in immediate recall of these items decreased, in general, from the top to the bottom of the matrix. It was also shown that the upper, left-hand corner of the matrices was in all cases the area of greatest accuracy or learning. A row-by-row decrease in accuracy from the top to the bottom of the matrices was quite marked when letter-matrices and word-matrices were used, while it was scarcely evident when symbols were employed. These symbol-matrices, on the whole, evoked more variability in spatial ordering of responses than any other type of material.

This note is concerned with an extension of the findings in Anderson and Ross's experiment. Specifically its purpose was (a) to determine the relationship between accuracy and temporal ordering of responses; (b) to discover the effect of a 5-min. delay before the final trial of 6 learning trials upon the accuracy of response; and (c) to obtain subjective reports from the Ss regarding the organizational processes employed in learning.

Materials. The stimulus-items used in this experiment were words and symbols; 25 of both kinds were used and they were identical to those employed by Anderson and Ross. These items were randomly arranged in separate, 5 × 5 matrices and printed on cards 8 in. square. There were nine different random arrangements of the word-matrices and symbol-matrices. Each arrangement of the word-matrices was carefully inspected to insure that adjacent words did not have strong associative value.

Nine Ss (7 men and 2 women) were used in the study. All were senior, undergraduate majors in psychology and students in the experimental class. Their ages ranged between 19 and 23 yr. They were arbitrarily divided into two groups: Group A of 5 Ss and Group B of 4 Ss.

[1] S. B. Anderson and Sherman Ross, Memory for items in a matrix, this JOURNAL, 68, 1955, 595-604.

The Ss in Group A were given symbol-matrix cards and those in Group B, word-matrix cards. Each S was given five 1-min. periods to learn them. Each trial was followed by a 30-sec. period during which S attempted to duplicate on a blank card the stimulus-card he had been studying. During these 30-sec. intervals, E, standing in back of S, recorded the temporal ordering of S's responses on another blank matrix-card. Following the first set of 5 trials, the Ss of both groups were given an irrelevant literary passage to read for 5 min. Then they were presented with another blank matrix and asked to recall as much of the stimulus-matrix as they could. Following this trial the types of stimulus-cards were switched between the two groups. Group A now learned the word-matrix and Group B the symbol-matrix. No random stimulus-arrangement was used more than once. All the Ss received a different random arrangement. This emphasis upon randomization was made to insure that cumulative results would be an index of *position-learning* rather than individual *item-learning*.

The 5-min. interlude between the second and third presentations, in which irrelevant literary passages were read, was omitted in the second half of the experiment since we expected the Ss to be aware of its ir-relevancy. Instead each S was given the following instructions: "Write just how you went about remembering the cards. That is, indicate the technique, if any, you used in memorizing or in recall."

To facilitate the presentation of results and discussion, the various cells in the matrix are indicated as follows: top row, reading from left to right, 1-5; second row, 6-10; third row, 11-15; etc.

Every cell was given an accuracy rank (1-25) on the basis of the total number of correct responses in that cell through all trials. Temporal order-ranks (1-25) were assigned on the basis of the mean temporal response-number, which were computed by dividing the sum of all the temporal response-numbers in any cell by the number of times the cell was filled in. That is, if a cell was filled in three times (first twice and fourth once), its mean temporal response-number would be $(1 + 1 + 4)/3$ or 2.0.

The rank-order correlation between 'total correct' and 'mean temporal response-number' for the 25 cells of the word-matrices was 0.90. Actually, five of these cells had *identical* 'total' and 'temporal' ranks. The same rank-order correlation for the symbol-matrices was 0.95, and the number of cells with identical 'total' and 'temporal' ranks was seven.

It is interesting to note Cells 2, 12, and 19 had identical 'total' and 'temporal' ranks, viz. 2, 10, and 25, respectively, in both the symbol-matrices and word-matrices.

The mean number correct ($N = 5$) on the fifth and sixth trials of the word-matrices was 21.0 and 21.56, respectively. A test of t showed no significant difference between means of correlated samples. The mean number correct ($N = 4$) on the fifth and sixth trials of the symbol-matrices was 17.0 and 19.8, respectively. Again there was no significance, although a t-value of 1.67 was obtained. We believe that further investigation along this line with a much larger N might reveal a significant reminiscence effect. More important, perhaps, is the present evidence that there is no indication of a *decrement* in the learning after a 5-min. period. In relation to this 5-min. recall-period, it is pertinent to note that interrogation of the Ss after the experiment showed that all but one S were totally unaware of the irrelevance of the reading period. In addition, none of the Ss reported that he made any attempt to recall the matrix during the reading period.

Inquiry of Ss as to how they learned the matrices revealed little. Only two Ss mentioned specifically in what order they memorized the materials. Both wrote that they had attempted to memorize a row at a time, beginning at the top and proceeding to the bottom of the matrix. Three Ss reported that they attempted to make sentences out of the words. Five Ss perceived that many of the symbols were in pairs, and they attempted to utilize this relationship in memorizing the matrices. This would indicate a possible defect in the materials used in this experiment. Since *item*-learning was not the important variable, and since *position*-learning was our primary concern, it might be better in subsequent studies to devise a set of symbols which do not permit a paired relationship.

The results of this study appear in general to substantiate the findings of the Anderson and Ross experiment. The following points of comparison may be made (1) The total correct responses in the symbol-matrices displayed the same emphasis on Rows 1 and 2. The emphasis on the lower, right-hand corner found on the earlier study was not apparent in this investigation. (2) The total correct responses in the word-matrices displayed the same spatial emphases as Anderson and Ross found. (3) The learning curves for each cell of the word-matrices display the same row-by-row increase in steepness as reported by Anderson and Ross. (4) The learning curves for each cell of the symbol-matrices showed some striking similarities with the earlier experiment. In both studies there was a row-effect in Columns 1 and 2, and a different row-effect in Columns 3, 4, and 5.

The following conclusions are proposed with respect to learning items

arranged in square matrix forms. (a) The temporal ordering of response correlates highly with accuracy of response in the various cells. (b) There is no significant difference between the total accuracy of responses in the final of five trials with the same matrix, and the accuracy of responses for one re-trial with this matrix after a period of 5 min. (c) Subjective evaluation by Ss of how they learned the matrices reveals little information with respect to any organizational processes which Ss may have employed.

University of Maryland   ALEXANDER W. ASTIN
                                          SHERMAN ROSS

AN EXPLANATION OF DU MAS'S "RADIAL ILLUSION"

In a recent number of this JOURNAL, Du Mas described an illusion, the "radial illusion," which he noticed one night while riding in a car in a rain that was falling straight down. "The raindrops," he reported, "appeared to be going in all directions (360°) from a point of radiation [*i.e.* the fixation-point]. I knew all the rain was falling down, but experientially some of the rain appeared to be going straight up."[1] . . . "It does not seem that physical factors are the cause of this phenomenon."[2]

Du Mas's observations are excellent—the illusion is compulsory under the conditions he described, and even more striking in a snow storm—but his denial of a physical explanation is hasty. It has, as the writer believes, the following physical basis.

The average-sized raindrop has a terminal velocity of about 26 ft. per sec. A car moving at 30 m.p.h. has a speed of 44 ft. per sec. The reader is invited to imagine or to sketch a simple diagram of the situation. Draw a base line equal to 44 units; at one end erect a perpendicular equal to 26 units. Complete the triangle with a hypotenuse. The hypotenuse represents the line of regard to the center of 'radiation' of the raindrops. Raindrops that lie along the hypotenuse have no angular motion relative to the eye: they are on what a Navy man would call a collision course with the eye. Drops above this angle move up relative to the eye, while those below move down. Dividing 26 by 44 yields the value of the tangent of the angle of the line of regard above the horizontal. Reference to a trigonometric table tells us that the angle above the horizontal is about 30° 35′; at 60 m.p.h., or 88 ft. per sec., the angle is reduced to about 16° 28′. A similarly simple computation on the basis of Du Mas's statement that the center of radiation was 12 ft. above the road and 12 ft. in front of him

[1] F. M. du Mas, The radial illusion, this JOURNAL, 69, 1956, 118.
[2] *Ibid.*, 120.

(and assuming that his eye was about 5 ft. above the road when he was seated in the car) shows that his line of regard was about 30° 15′ above the horizontal. The similarity to the calculated value of 30° 35′ is remarkable, and probably coincidental. Du Mas's raindrops may not have been falling at the theoretical average speed, or he may not have been going at exactly 30 m.p.h. when he made the observation. He failed to observe, or at least to report, that the height of the 'center' diminishes as the speed of the car increases—which it clearly does.

Du Mas actually approached this answer in his note when he cited the Gibson paper.[3] His objection to that explanation was that the ground is stable and rain is not—but relative motion is relative motion however come by.

Central Institute for the Deaf   BRUCE H. DEATHERAGE
St. Louis, Missouri

A FORCED-CHOICE METHOD OF LIMITS

The question of the validity of psychophysical measures, raised recently by Blackwell for visual judgments, is especially persistent for chemoreception.[1] In research on both taste and smell it has long been considered desirable to avoid the adapting effects of strong stimuli by using what may be called an "undirectional method of limits," that is, a method of limits consisting of ascending series only. Such a method obviously involves serious problems concerning the influence of self-instructions by O, and the fact that very *reliable* results may be obtained does not fully solve these problems.[2] On the other hand, especially when one is concerned with collecting data from fairly large numbers of Os, the method of limits has considerable appeal as a means of saving time.

A simple revision of procedure for the method of limits eliminates the problem of differences in criterion among the Os, and also eliminates the errors of anticipation and habituation. This revision consists of forcing O to indicate which of four stimulus-presentations, at each ascending level, actually contains the taste or odor to be detected. Three of the four presentations are 'blanks,' while the other contains the stimulus-material of interest. This method has recently been used by Comrey with considerable

[3] J. J. Gibson, Paul Olum, and Frank Rosenblatt, Parallax and perspective during aircraft landings, this JOURNAL, 68, 1955, 372-385.

[1] H. R. Blackwell, Psychophysical thresholds, *Engin. Res. Bull.* No. 36, University of Michigan, 1953, 1-227.

[2] F. N. Jones, The reliability of olfactory thresholds obtained by sniffing, this JOURNAL, 68, 1955, 289-290.

success for a type of judgment which would have caused difficulty with a simple method of limits.[3]

In actual practice one first prepares a series of stimulus-materials in ascending order of magnitude, just as for the usual method of limits. When the series is presented, however, three blanks are presented with the series-stimulus, the order of the four for each level being predetermined from a table of random numbers. Thus Comrey offered O four "sniff bottles" at each level, only one of which contained odorous material. The sniffs could have been 'blasts,' or, for taste, swigs. Since O has one chance in four of guessing correctly, E may wish to require two or more successive correct choices, the probability of two successive successes being, of course, $1/16$, of three $1/64$, and so on. Or two or more complete series may be used. The chance probability which E is willing to tolerate determines the final decision. By using one comparison-stimulus and three standard stimuli the method may be adapted to the determination of difference thresholds. Also, either more or fewer blanks may be used, with consequent alteration of the probabilities.

If we take as our basic criterion of the validity of a psychophysical method the requirement that thresholds obtained by the given method should depend primarily upon events in the sensory system under consideration and upon as little else as possible, the forced-choice method of limits has at least face validity. It will not, of course, cure all ills of poor stimulus-control, but it does eliminate differences in criterion from O to O, as well as the more usual pitfalls of the method of limits.

University of California, Los Angeles F. NOWELL JONES

PSYCHOLOGIE FRANÇAISE

The first number of *Psychologie Française*, a quarterly journal published by the French Society of Psychology, has been received. It was founded with the purpose of reporting all the activities of the Society: of its general monthly meetings as well as of the meetings of its various sections (clinical, child and education, psychophysiological, social and applied). It is, however, more than a mere bulletin of the Society's activities as it publishes in addition many short scientific articles and original researches.

The first number opens with an editorial by P. Fraisse, the secretary general of the Society, in which the objects and aims of the new journal

---

[3] A. L. Comrey, Elizabeth Klein, and G. H. Watson, Sensitivity to the odor of Vitamin B_1 (unpublished study).

are outlined. This is followed by 12 short articles by various members of the Society, as follows: Données récentes sur les mécanismes de régulation de l'appetit, by J. Le Magnen; Perception et fixation mnemonique, by C. Flores; Étude des phénoménes d'enrayement et d'activation du comportement par stimulation thalamique et réticulaire chez le rat non anesthésié, by V. Bloch and D. O. Hebb; Contribution au diagnostic de l'épilepsie par les moyens de l'analyse graphométrique, by R. Perron and H. de Gobineau; Recherches sur la nature des insuffisances symboliques dans les retards du language et dans les débilités intellectuelles, by N. Galifret-Granjon; Influence des changements de réseaux de communication sur les performances des groupes, by C. Flament; Les relations affectives vues a travers le mensonge, by M. Durandin; Sur l'analyse hiérarchique, application et contribution a une techbique de construction d'échelles d'attitude, by S. Moscovici; Réflexions sur le délire, by H. Ey; La formation professionelle des ouvriers spécialises, by M. Leplat; and La méthode de l'étude psychologique des altitudes, by E. Deutsch.

The number closes with the proceedings of the meetings of the Society and of its various sections. The first number, a 24-page octavo, is printed on eggshell paper with the original articles in large type and the proceedings of the meetings in small.

The JOURNAL welcomes the new periodical into the psychological fold.

K.M.D.

SIXTY-FOURTH ANNUAL CONVENTION OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

Approximately 4,500 psychologists and friends registered for the Sixty-Fourth Annual Convention of the American Psychological Association in Chicago August 30-September 5, 1956. A total of 556 hours were officially set aside for the reading of scientific papers, for symposia, and for special papers on a variety of different subjects.

Every annual convention is highlighted by its own particular events. This convention was marked by the first award of the Association's "Gold Medal for Distinguished Service to Psychology," to Robert Sessions Woodworth, Professor Emeritus of Psychology, Columbia University, and of awards for "Distinguished Scientific Contributions" to Professors Wolfgang Köhler, Kenneth W. Spence, and Carl R. Rogers. These awards are hereafter to be made annually in recognition of meritorious service to psychology.

Among the special papers attracting attention were Theodore New-

comb's presidential address on "The prediction of interpersonal attraction"; Gardner Murphy's address, in commemoration of Freud's centennial, on "The current impact of Freud on American psychology"; and the addresses of two guest speakers: Hans Selye, Director of the Institute of Experimental Medicine and Surgery at the University of Montreal, on "The psychosomatic implications of the general adaptation syndrome," and J. R. Reese, Director of the World Federation for Mental Health, on "International aspects of mental health."

The sectional sessions were well attended and many caught the attention of the press as well as of psychologists. The extensive press coverage indicated the interest with which the public as a whole is viewing developments in psychology.

At the annual Business Meeting, Lee J. Cronbach was inducted as President of the Association and Harry F. Harlow as President-Elect, and the following elections were announced: of Anne Anastasi and Harold M. Hildreth to three-year terms on the Board of Directors; of Norman L. Munn, to the editorship of *Psychological Monographs* for the term 1958-63, and of Allen J. Sprow to the executive editorship of the *Psychological Abstracts*. The elections of 1,269 Associates and 115 Fellows were also announced, bringing the membership of the Association to 15,545, among whom there are 1,873 Fellows.

Among other actions taken by the Council of Representatives were the following: the establishment of a Division of Engineering Psychology; a decision to publish longer and fewer *Psychological Monographs* which emphasize full reports of fundamental research; the appointment of an *Ad Hoc* Planning Group on the Role of the Association in Mental Health Programs and Research; the approval for organizing a conference on experimental design and inference in psychology; a decision to publish an annual biographical directory of the Association's membership; the purchase of the *Journal of Educational Psychology;* and the adoption of policy statements concerning social issues, dissemination of information, utilization of psychologists in public service programs, and the handling of requests for funds from foundations and other sources.

The 1957 meeting will be held in New York City.

Washington, D.C.   ROGER W. RUSSELL

Samuel Weiller Fernberger: 1887-1956

Samuel Weiller Fernberger died of acute leukemia on May 2, 1956, in Philadelphia. He was born in that city on June 4, 1887, was graduated from the William Penn Charter School, and received his B.S. in 1908, his M.A. in 1909, and his Ph.D. in 1912 from the University of Pennsylvania. While a graduate student, he held the positions of assistant and, later, instructor, in psychology at that University. In 1912 he went to Clark University as instructor, and in 1915 he was made assistant professor. He returned to his alma mater in 1920 at that rank, became professor in 1927, and was made emeritus professor in 1956. He was married to Eve Wallerstein in 1922 and is survived by her and their two sons, Edward and John.

For many years Fernberger was active in the affairs of psychological associations. He was treasurer of the American Psychological Association from 1922 to 1924, secretary from 1926 to 1928, and president of the Division of General Psychology in 1952. He was president of the Eastern Psychological Association in 1936, secretary of the Society of Experimental Psychologists from 1928 to 1948, and president of the National Institute of Psychology from 1936 to 1941. He was editor of the *Psychological Bulletin* from 1918 to 1930, editor of the *Journal of Experimental Psychology* from 1930 to 1946, and coöperating editor of *The American Journal of Psychology* from 1925 to his death.

Fernberger served his country with great energy and distinction in both World Wars. During 1918 and 1919 he was second lieutenant and then first lieutenant in the infantry of the United States Army, and he received the Croix de Guerre and the Bronze Star for action in France. From 1940 to 1945 he was active in the National Research Council and the National Defense Research Council in various capacities, working on problems of perception, training, and human engineering. He was awarded the President's Certificate of Merit for these services.

Fernberger's publications are impressive for both their number and the variety of interests that they represent. He was first and last a psychophysicist, but he was able to make contributions to many other fields. Within the area of psychophysics, he dealt with such problems as the introspective properties of psychophysical judgments, intra-serial effects, the status of the middle category of judgment, effects of practice, the 'absolute method,' the effects of physical and mental work, the stimulus-error, judgment-times for

the various categories of judgment, and many others.[1] A number of these papers are classics in their fields.

Outside of psychophysical methodology, his contributions on the facial expression of emotion, on memory-span, and on the range of visual apprehension[2] are frequently cited. He also wrote on such varied matters as apparent movement, imagery, aphasia, the history of psychology, experiences resulting from the taking of peyote, and the psychological interpretation of Sioux Indian shields.[3] The last of these represents one aspect of a long-standing interest in anthropology; among his collections was one of aboriginal masks. His one book, *Elementary General Psychology*,[4] published in 1936, shows the pronounced influence that Titchener had upon him.

At all stages of his career he exhibited an interest in those applications of experimental psychology that employed rigorous, preferably psychophysical, methods. He once testified before the Public Service Commission of Pennsylvania as an expert for the complainant in an action brought by the Yellow Cab Company of Philadelphia to prevent the use by others of taxicabs with colors alleged to resemble too closely the color of Yellow Cabs. His testimony involved not only verbal opinions but also the performance of demonstrations in the hearing room in the presence of the examiner and the interested parties. He mixed colors on a rotary color mixer to match the colors of the complainant's and respondent's cabs, showed that they were closely similar—they differed by not more than three just notice-

[1] *E.g.* An introspective analysis of the process of comparing, *Psychol. Monogr.*, 26, 1919 (No. 117), 1-161; Interdependence of judgments within the series for the method of constant stimuli, *J. exp. Psychol.*, 3, 1920, 126-150; Die Ungewissheitsurteile in der Psychophysik, *Arch. ges. Psychol.*, 80, 1931, 273-290; The effects of practice in its initial stages in lifted weight experiments and its bearing upon anthropometric measurements, this JOURNAL, 27, 1916, 261-272; On absolute and relative judgments in lifted-weight experiments, *ibid.*, 43, 1931, 560-578; The influence of mental and physical work on the formation of judgments in lifted weight experiments, *J. exp. Psychol.*, 1, 1916, 508-532; An experimental study of the "stimulus error," *ibid.*, 4, 1921, 63-76; (with E. Glass, I. Hoffman, and M. Willig) Judgment times of different psychophysical categories, *ibid.*, 17, 1934, 286-293.

[2] *E.g.* A preliminary study of the range of visual apprehension, this JOURNAL, 32, 1921, 121-123; (with Ellen Jarden) The effect of suggestion on the judgment of facial expression of emotion, *ibid.*, 37, 1926, 565-570; (with P. M. Martin) Improvement in memory-span, *ibid.*, 41, 1929, 91-94.

[3] *E.g.* Fundamental categories as determiners of psychological systems: an excursion into ancient Chinese psychologies, *Psychol. Rev.*, 42, 1935, 544-554; New phenomena of apparent visual movement, this JOURNAL, 46, 1934, 309-314; Possible effects of the imaginal type of the subject on aphasic disturbances, *ibid.*, 30, 1919, 327-336; Observations on taking peyote (*Anhalonium Lewinii*), *ibid.*, 34, 1923, 267-270; (with Frank G. Speck) Two Sioux shields and their psychological interpretation, *J. abnorm. soc. Psychol.*, 33, 1938, 168-178.

[4] Reviewed in this JOURNAL, 49, 1937, 490 f.

able differences along the base of a model of the color solid that was also produced—and that their resemblance increased as the illumination of the room was decreased.

He also enjoyed making statistical studies that might be thought of as 'footnotes to history.' These appeared at intervals from as early as 1917 to the last year of his life. They treated such topics as the number of articles in psychology published in different languages, an analysis of members of A.P.A. as to geographical distribution, origin of Ph.D., and positions and research interests of men and women psychologists, the publications of American psychologists, the history of A.P.A., Wundt's students, and the like.[5]

As a psychologist, Fernberger was concerned more with concrete fact and the details of method than with general theory. He was fond of insisting to his students, with a vigor that bordered upon paradox, that results are a function of the method used in obtaining them. His dissertation,[6] carried out under the direction of F. M. Urban, compared the methods of constant stimuli and just perceptible differences, somewhat to the advantage of the former; this set the tone of a great deal of his later work. Perhaps he was never more in his element than when he was acting as an observer or experimenter with lifted weights on a rotating table. His feeling for the method of constant stimuli seemed to amount to a positive affection. Guilford's demonstration of very high correlations between linear interpolation and the constant process never modified his attachment to the Müller-Urban treatment. The writer remembers with what trepidation he and Leon Arons once had to confess to Fernberger that (to study contrast effects) we had put nothing but 100-gm. weights on the table during an experimental session in which he had made a long series of judgments with his usual confidence. We need not have worried, since he regarded the incident as a huge joke which he long enjoyed recounting to others at his own expense.

He was a man of intense loyalties and, if he was conservative in his attitudes toward psychology, it is difficult not to attribute this in good part to the effects of the influences exerted upon him by Urban during his

---

[5] E.g. On the number of articles of psychological interest published in the different languages, this JOURNAL, 28, 1917, 141-150; Statistical analysis of the members and associates of the American Psychological Association, Inc. in 1928, Psychol. Rev., 35, 1928, 447-465; The publications of American psychologists, ibid., 37, 1930, 526-544; The American Psychological Association: A historical summary, 1892-1930, Psychol. Bull., 29, 1932, 1-89; Wundt's doctorate students, ibid., 30, 1933, 80-83; On the number of articles of psychological interest published in the different languages: 1946-55, this JOURNAL, 69, 1956, 304-309.
[6] On the relation of the methods of just perceptible differences and constant stimuli, Psychol. Monogr., 14, 1913, (No. 61), 1-81.

doctoral work at Pennsylvania and by John Wallace Baird and Titchener in his years at Clark. He was invited to Clark by Baird, a disciple of Titchener, whose main theme was "systematic experimental introspection." In Fernberger's own words, "[Titchener] considered Baird one of his intelectual sons, and when Baird died, for some reason (I suppose it was because I was still at Clark), he took me on as a sort of grandson."[7] A taste of the atmosphere, which was to have a permanent effect upon him, is provided by the following anecdote. Professor E. G. Boring writes, "I remember so well the 1913 meeting of the APA in New Haven when Baird undertook to demonstrate experimental introspection to the group. Clark graduate students came in, were presented with some sort of stimulus, and gave their introspections. Baird was all enthusiasm and felt sure it would convince people of the value of the method, but actually it was a fiasco. Still, the group kept their self-assurance, and that included Sam."[8] To a member of the next generation, this fixture upon a losing cause seems hardly the most favorable preparation for the actual turn of events in psychology after 1913; but the cannonading of the behavioristic revolution was seldom heard in the lifted-weight room.

Fernberger was a fascinating and successful teacher. For many years his Thursday afternoon graduate seminar on Problems of Psychology was a principal focus of intellectual excitement in the department at Pennsylvania. On various memorable occasions, the roles of James, Wundt, McDougall, Titchener, Watson, Köhler, Freud, and others were dealt out to the members of the seminar, who were obliged to assume them in the discussion of a variety of questions. Fernberger used to claim with great glee that the students took on not only the theoretical attitudes of their noted exemplars but their personal traits as well. He was equally effective with undergraduates, and, to his last semester of teaching, students often infringed upon regulations by attending his lectures in elementary psychology instead of those to which they had been assigned.

He was a staunch exponent of individual and academic freedom. In this connection, he was the subject of some newspaper notoriety when, in 1954, he resigned from the Philadelphia Chapter of the Sons of the American Revolution because the Chapter planned to give honor to Senator Joseph R. McCarthy, whose methods he felt were repugnant to the ideals that both he and his ancestors had fought for.

Fernberger planned to retire in June, 1956. At Easter time he fell ill,

[7] Prepared remarks, referred to below.
[8] Personal communication. The following evaluation is strictly the writer's.

however, and he never fully recovered. Plans had been made for a festive occasion on May 4, a time at which he was to have been honored by a dinner, the presentation of a bound volume of letters written to him by his colleagues and his doctoral students, and the unveiling of a bronze tablet on the wall of the Samuel W. Fernberger Library of psychological books and journals that, as a life-time collection, he had contributed to the University of Pennsylvania. He knew of these plans and had, indeed, prepared remarks for the occasion. The grief of his friends was sharpened by the sadly ironic circumstance that his funeral occurred on the day that had been chosen for the celebration. Excerpts from his modest words of thanks and self-evaluation were read at the funeral. They mentioned many of his friends and teachers, among whom Urban, Baird, and Titchener figured prominently. Those who survive him have some consolation in the fact that he had known how widely and well his own strong loyalties and friendships were returned.

University of Pennsylvania   FRANCIS W. IRWIN

Lightner Witmer: 1867-1956

With the death of Lightner Witmer on July 19, 1956, at the age of 89, an era in American psychology has passed. Lightner Witmer was the last surviving charter member of the American Psychological Association. On this count alone, a loss must be felt on the psychological horizon, but his career does much more than serve as a reminder of the rapidly changing scene that is American psychology.

Dr. Witmer founded the Psychological Clinic at the University of Pennsylvania in 1896. This was the first psychological clinic in the United States and, indeed, in the world. Its influence helped to lead to the founding of many other clinics in the universities and school systems throughout the country. His importance in helping to establish this work is epitomized in the fact that he was the first to use the term, *clinical psychology* for this area of endeavor.

Born in Philadelphia, on June 28, 1867, Witmer attended the Episcopal Academy and graduated from the University of Pennsylvania in 1888. George S. Fullerton, who had brought James McKeen Cattell to the university, asked him to become an assistant in psychology to Cattell, a post which he accepted. Undoubtedly part of his later interests could be attributed to this period of study with Cattell, that doughty champion of the heretofore neglected topic of individual differences. Witmer, himself,

asserted Cattell was more influential upon his thinking and development than was any other man with whom he studied. Subsequently, he went to Leipzig where he worked under Wilhelm Wundt, receiving his doctorate there in 1892. While at Leipzig he was a member of a distinguished company that included Edward Bradford Titchener, Oswald Külpe, Frank Angell, Lincoln Steffens, and Hugo Eckener. On returning to this country he was made Director of the Laboratory of Psychology in the University of Pennsylvania, for Cattell had accepted appointment at Columbia the previous year. In subsequent years for brief periods of time, he also served on the faculties of Bryn Mawr College and Lehigh University. At both of these institutions he established psychological laboratories. In 1907 he founded the journal *Psychological Clinic* in which for twenty years his principal contributions to the psychological literature appeared. On his retirement in 1937 he had served on the faculty of the University of Pennsylvania for forty-five years.

He was little known as a person to most psychologists of today. His contemporaries regarded him as aggressively honest and critical toward others but equally ready to criticize and correct his own opinions. He was a clear thinker and his expression was specific and to the point, as examination of his writings reveal.

As were his teachers, Wundt and Cattell, before him, Lightner Witmer was a pioneer. He was in the forefront of those who would put psychology to work for the good of their fellow-men. His trail-blazing task was carried on in one major region of the application of psychology. From the challenge in March 1896 of a case of a child who could not learn to spell to the end of his career, he was interested in the area of diagnostic and remedial work with the intellectually and educationally deviating child. Although conceiving of pure and applied psychology as advancing on a 'single front,' he steadfastly worked in the area this first school child identified for him.

Leaders in specific fields, who were first his students and later his colleagues, found their places in the expanding operation of the Psychological Clinic. Edwin B. Twitmeyer specialized in the diagnosis and correction of speech defects, Morris S. Viteles in vocational and industrial guidance, and Robert A. Brotemarkle in the personality adjustment of college students.

Witmer pioneered in collaborative clinical work with representatives of other disciplines. Even during the first year of his Clinic's operation, he sought the help of a teacher of the deaf with one of the children who came to him. In this and subsequent collaborative efforts with representatives of

other disciplines, principally teachers, social workers, and neurologists, the team approach was firmly established; but the teacher was always central to this collaboration.

In his clinical work attention to the quantitative results of psychometrics was parallelled by the use of the qualitative findings of his other diagnostic procedures. Remedial plans and subsequent treatment were also essential ingredients of his approach.

The applied emphasis, the centering on the problems of the individual, the collaborative approach of specialists to the task at hand, the mingling of quantitative and qualitative findings about the person concerned, and the use of diagnosis and treatment as integral parts of the process, are all typical aspects of clinical psychology as we now know it.

Two significant facets of clinical psychology, as we know it today, were, however, lacking then. The 'dynamic approach,' to use a convenient short-hand label, was almost completely ignored. He was little influenced by the French psychiatric thought of his day and not at all by Freud. The second missing piece from today's crazy-quilt that is clinical psychology arises from essentially the same omission. This was the lack of clinical collaboration with 'analytically orientated' psychiatrists. It was not part of his experience to work closely with the pioneers of modern psychiatry. Instead, he turned to men who were primarily neurologists in working out the medical phase of his professional collaboration.

As a result his major influence embodies a paradox. He had less influence upon that body of workers today called *clinical psychologists,* who in considerable number adopt a dynamic approach and work in a psychiatric setting, than upon other psychologists. He was most influential with the workers in those other fields that use the method of clinical psychology but name their areas of endeavor differently. The lineal descendents of Lightner Witmer are found in the far-reaching applications of basic clinical psychology to individual problems in education, in vocation and industry, in speech correction, in socio-individual adjustment, a span of interests affecting a wide range of human behavior.

Northwestern University

ROBERT I. WATSON

Chauncey McKinley Louttit: 1901–1956

Chauncey McKinley Louttit, editor of the *Psychological Abstracts* and chairman of the Department of Psychology at Wayne University, Detroit, Michigan, died of acute leukemia in his fifty-fifth year on May 24, 1956,

at his home in Ferndale, Michigan. He was born in Buffalo, New York, on October 9, 1901.

Following his elementary and secondary education in the public schools of Buffalo, he went to Hobart College, Geneva, New York, from which he was graduated in 1925 with a B.S. degree. Then, after a research fellowship at the Training School, Vineland, New Jersey, he entered the graduate school of Yale University where he worked with Professors Raymond Dodge and R. M. Yerkes on the reflexes[1] and reproductive behavior of the guinea pig,[2] respectively and from which he received his Ph.D. degree in 1928.

Immediately after receiving this degree, Dr. Louttit went to the University of Hawaii as research associate in its psychological clinic. He remained there two years, from 1928-1930, during the second of which he served in addition to his other duties, as psychologist in the Kamehameha Schools of Honolulu. In the fall of 1930, he went to Ohio University, Athens, Ohio, as an assistant professor. After one year in this position, he accepted a call from Indiana University as assistant professor and director of its psychological clinic, a position he held until 1940. During his service there he published his book, *Clinical Psychology: A Handbook of Children's Behavior Problems*, in 1936, and was promoted to an associate professorship in 1938.[2a]

In 1940, on leave of absence from Indiana University, Dr. Louttit began his service with the U. S. Navy which lasted for the duration of World War II. He served successively as clinical psychologists at the U. S. Naval Medical School in 1940-1941; as assistant chief of the Psychological Division in the Office of the Coördinator of Information in 1941-1942; as assistant officer in charge of the Quality Control Section, Training Division, Bureau of Naval Personnel in 1942-1944; as commanding officer of the Naval Training School at Camp Macdonough from February to October of 1944; and as commanding officer of the Service School Command, Naval Training Center, Bainbridge, Maryland, in 1944-1945. He entered the Navy as a lieutenant commander; when he was separated from the Service, he was a captain.

---

[1] Raymond Dodge and C. M. Louttit, Modification of the pattern of the guinea pig's reflex response to noise, *J. comp. Psychol.*, 6, 1926, 267-285.
[2] Louttit, Reproductive behavior of the guinea pig: I. The normal mating behavior, *ibid.*, 7, 1927, 247-263; II. The ontogenesis of the reproductive behavior pattern, *ibid.*, 9, 1929, 293-304; III. Modification of the behavior pattern, *ibid.*, 9, 1929, 304-315.
[2a] *Op. cit.*, 1-695. He published a revised edition of this book in 1947 and was arranging for the publication of a second revised edition at the time of his death.

In the fall of 1945, Dr. Louttit went to Ohio State University as professor of psychology and director of the psychological clinic. He served there one year and then, in 1946, went to Sampson College as dean of faculty. This institution, a temporary junior college, had just been established by the State of New York, at the Sampson Naval Training Base near Geneva, to meet the demands of the returning service men for a college education. Before the compltion of his first year there, he accepted, in the spring of 1947, the position of executive dean of the Galesburg undergraduate Division of the University of Illinois—a temporary junior college established for the returning service men in Illinois. When that Division was closed in 1949, he was transferred to the University of Illinois as assistant to the Provost. He retained this position until the fall of 1954 when he accepted the call to a professorship at Wayne University —to the position he held at the time of his death.

Dr. Louttit assumed the editorial responsibility of the *Psychological Abstracts* in January 1947, while at Sampson College.[3] He succeeded Professor W. S. Hunter, who had edited the *Abstracts* since its founding in 1927. He had a difficult task and moreover, Hunter, because of his high standards and familiarity with the work, was a difficult man to follow. Louttit, however, picked up the work and carried it with distinction through ten volumes (Vols. 21-29) despite his heavy administrative duties and changes in his positions.

Louttit's interest in the literary aspects of psychology and in editorial work showed itself early in his career. His first publication, which appeared while he was an undergraduate at Hobart College, was "A bibliography of sex differences in mental traits."[4] This was followed, while he was a graduate student at Yale University, by a "Bibliography of bibliographies: 1900-1927,"[5] in which he listed and gave a subject matter index of 2134 bibliographies in psychology, and by an article on "The use of bibliographies in psychology,"[6] in which he evaluated the types of bibliographies found in psychology and discussed the desiderata involved in their collection and publication.

These early publications were followed, through the years, by many other

---

[3] W. S. Hunter, Editorial note, *Psychol. Abstr.* 20, 1946, 469; C. M. Louttit, Editorial note, *ibid.*, 21, 1947, 1.
[4] *Op. cit., Tr. School Bull.*, 22, 1925, 129-138.
[5] *Op. cit., Bull. Nat. Res. Coun.*, 1928 (No. 65), 1-108.
[6] *Op. cit., Psychol. Rev.*, 36, 1929, 341-349.

articles from his pen on editorial matters.[7] In 1937 he undertook his first editorial task when he assumed the business editorship of *The Psychological Record*—a new journal founded by the Principia Press, Inc., of Bloomington, Indiana, "to afford authors the opportunity of immediate publication."[8] In addition to this duty, which he carried until December 1942, when the new journal was suspended because of the war effort, he edited two issues of the *Directory of Applied Psychology;* the first in 1941 and the second in 1943. His editorial duties were not, however, laid down for long; soon after the War, as we have seen, he took over the editorship of the *Abstracts.* In the discharge of this duty, he frequently found it desirable to publish Notes about matters of editorial concern. In 1952, he was chairman of the Committee of the Council of APA Editors that prepared a "Publication manual" for the use of contributors to the journals of the American Psychological Association.[9]

His interest in administration also showed itself early in his career. He always, as will be recalled, left purely academic positions, when opportunity availed, for administrative positions even though they were nothing more than the directorships of psychological clinics. His flair for administration was fully tested and developed, however, in the Navy. He discovered his forte while in the Services and all the positions held by him after the War were administrative.

Dr. Louttit's chief interest in psychology was in the applied fields—in particular in clinical psychology, in which he was a diplomate. In recognition of his efforts and accomplishments in the applied fields, his confrères elected him to the presidency of the American Association for Applied Psychology in 1943.

In the death of Dr. Louttit, psychology has lost an outstanding editor and an ardent worker in the applied field. The rare combination of talents that existed within him shall not soon be seen again.

He is survived by his wife, Laura Talcott Louttit, and two sons, Robert Irving and Richard Talcott Louttit.

KARL M. DALLENBACH

University of Texas

---

[7] Louttit, Psychological journals: A minor contribution to the history of psychology, *Psychol. Rev.,* 38, 1931, 455-460; *Handbook of Psychological Literature,* 1932, 1-273; A study of students' knowledge in the use of the library, *J. appl. Psychol.,* 16, 1932, 475-485; *A Proposed Decimal Classification for Psychological Literature,* 1932, 1-30; The Dewey decimal system and psychology, *J. gen. Psychol.,* 9, 1933, 234-238; Psychological journals, this JOURNAL, 46, 1934, 147-148; Publication media available for applied psychology, *J. appl. Psychol.,* 24, 1940, 85-91; Library classification for psychological literature, *Psychol. Rec.,* 4, 1941, 350-364.

[8] Editorial announcement, *Psychol. Rec.,* 1, 1937, 2.

[9] *Op. cit., Psychol. Bull.,* 49, 1952, 389-449.

James Pertice Porter: 1873-1956

James P. Porter, the son of Alfred and Elizabeth (Marksburg) Porter, was born at Hillsboro, Indiana, on September 23, 1873. He died as the result of a heart condition at Swarthmore, Pa., at the home of a daughter, on September 15, 1956, shortly before his eighty-third birthday.

An internationally known psychologist, Dr. Porter received his A.B. degree at Indiana University in 1898, his A.M. at Indiana in 1901, and his Ph.D. at Clark University under G. Stanley Hall in 1905. In 1917 he was awarded an honorary Sc.D. by Waynesburg College.

Like many others of his day, Porter began his teaching career when a mere boy, in 1891 at age 17 years, in the high school at Newtown, Indiana, after he had attended the Indiana State Normal School for one year. Later he served first as a high school principal in Illinois and then as an instructor in psychology at Indiana University. Subsequently he taught psychology at Clark where he became Dean of Clark College prior to his coming to Ohio University as head of the Department of Psychology in 1922. He was named Professor of Psychology Emeritus at Ohio University in 1943.

Porter rendered distinguished service to psychology as a profession during a period of its most rapid growth and expansion as a science. He is perhaps best known as editor of the *Journal of Applied Psychology,* a periodical which he owned and edited from 1920 to 1943. At the time of his retirement he arranged for the preservation of this journal by transferring its ownership and management to the American Psychological Association. To him the assurance that his journal would not die was a matter of the utmost importance.

He was noted for his calmness, warmth, affability and kindliness. If he could be classified, he would have to be catalogued as an extrovert. He had many friends and belonged to many social and professional organizations; but he was no mere joiner for he contributed to every organization that he joined and worked hard for its welfare.

He was a member of Sigma Chi, Phi Beta Kappa, the Ohio Teachers Association, the National Education Association, the Society for the Advancement of Education, the Ohio College Association (President of the Section on Psychology in 1933), the Ohio Academy of Science (President in 1935), the American Association of Applied Psychologists, the Midwestern Psychological Association (President in 1941-1942), and the American Psychological Association. He was also a member of the International Congress of Zoölogy, a delegate to the International Zoölogical Congress, the International Congress of Psychology and Psychotechnique, and the International Congress on Eugenics.

In 1918 Dr. Porter served as a captain in the psychological service of the U. S. Army and at the conclusion of World War I became a major in the reserve. In 1944-1945 he served in the office of the Adjutant General as a counselor to returning soldiers. When the veterans flooded the college campuses following World War II, he became an instructor in psychology in the Extension Division of the University of Illinois at Danville.

Although Porter had wide interests, he was concerned most of all in the application of psychological principles to the practical affairs of everyday life. During his early days he made special studies of the English sparrow, of spiders, and of intelligence and imitation in birds. In his later years he was keenly interested in intelligence tests for high school and college students. He was one of the first to recognize the value of these tests as one means of helping to identify those students who might well work for advanced degrees, and it was his custom to encourage these students to major in psychology. He often loaded several of these students in his auto and took them with him to professional meetings, thus enabling them to meet and to become acquainted with the leaders in the field. In this way he inspired a number of able young men to enter the profession of psychology. A study made in the central office of the American Psychological Association to find where members of the Association had most often received their undergraduate training revealed that, in proportion to its enrollment, Ohio University had produced more members than could be accounted for on the basis of chance factors.

Dr. Porter had so many friends, drawn from acquaintances and his onetime students, that there are now many to miss him and regret his passing. He was a man who taught his students not only subject matter but also, by his example, the art of living.

Ohio University   HARVEY C. LEHMAN

BOOK REVIEWS

Edited by M. E. Bitterman, The Institute for Advanced Study

Fundamentals of Language. By Roman Jakobson and Morris Halle. The Hague, Mouton and Company, 1956. Pp. ix, 87.

Psychologists and linguists have a lot in common, more than appears on the surface. Both attempt to analyze behavior; both search for uits of analysis and methods of recognizing those units; both are concerned lest their analysis should destroy the delicate texture of the behavioral configuration that gives meaning to the separate units; both try to analyze perception; both search for attributes of the stimulus capable of supporting discrimination; and both have made conscientious efforts to be objective in dealing with their problems. The linguist is concerned only with a particular segment of the total spectrum of behavior and perception, but within this segment his problems are little different from those of the psychologist. In fact, this very restriction of interest to verbal behavior has in many respects proved an advantage to the linguist, for he has been able to refine his concepts and sharpen his definitions to a degree that psychologists might properly envy.

Take, for example, the definition of the *phoneme*. Here is a topic that has inspired controversy among linguists for almost a century. The amount of heated disagreement at first persuades the visiting psychologist that linguists are confirmed quibblers incapable of settling their most fundamental questions in any satisfactory way, but, after some reflection on the sad state of the phoneme, a psychologist is apt to remember his own difficulties with the definition of *response*. A phoneme is, after all, a response, in just the sense that psychologists use the term, and the linguist's problems of definition bear a striking resemblance to our own. Of course, a phoneme is also a *stimulus*, but this alternative is scarcely more attractive to the psychologist, who has seldom defined stimulus better than he has defined response.

Psychologists have managed to proceed in spite of their semantic disagreements because they can usually settle the question for each particular experiment. In a similar way, linguists who hold very disparate views on the nature of a phoneme can usually agree on the set of phonemes to be used in analyzing any particular language. Behind the practical applications, however, there always lurks a suspicion that the experiment, or the language, would have seemed simpler, more intelligible, more closely tied to the rest of the science if the right definition of the analytic unit had been used.

Jakobson and Halle have a firm conviction as to what is the right way to define linguistic units, and the first half of this book is an exposition of the merits of their approach. In passing, however, they review and reject a number of alternatives. In order to see the similarity between the problem of defining the phoneme and the problem of defining the response, it is an interesting exercise to translate these alternatives from linguistics to psychology. Let us see how the same arguments would sound if a psychologist made them.

"There is an 'inner' and an 'outer' way to define response," the psychologist might say, "depending upon whether the definitional criteria are considered to be intrinsic in the behavior itself or to reside in some external realm to which the behavior is allegedly related. The 'outer' approaches to the response can be classified as follows:

"*A*. The *mentalist* view, which is the oldest still surviving, holds that the response is an act imagined or intended, opposed to the actual physical movements as a psychological phenomenon to the physiological fact. It is the mental equivalent of an exteriorized movement.

"*B*. The *code-restricting* view attempts to locate the response in a socially adopted code, and assigns the variant forms of the response to individual differences in behavior.

"*C*. The *generic* view appeals to the notion of resemblance. A response is a family or class of movements related by the fact that they appear similar to an observer.

"*D*. The *fictionalist* view is that a response, like all scientific concepts, is nothing more than a convenient myth that psychologists maintain in order to facilitate theory. The response is a theoretical fiction that does not exist in the particular, discrete movements that comprise our actual behavior.

"*E*. The *algebraic* view would substitute algebraic symbols for movements and then concentrate upon an analysis of the sequences of symbols. Two different symbols which never occur in the same context can be grouped together into a single response, since no confusion will result from calling them both by the same name. Thus the definition of the response is referred to the context, rather than to its empirical correlates.

"Opposed to these 'outer' approaches, which refer the response to something other than the actual movements (to the mind, a social code, an observer, a theory, the context), is the 'inner' approach, which attempts to locate the critical dimensions of the response within the objective movements themselves. Movements may be characterized as weak or powerful, short or sustained, vocal or motoric, total or segmental, upward or downward, etc. These are distinctive features of responses. A response is simply a bundle of these features."

The translation is not entirely plausible, since the algebraic view is peculiarly linguistic with no real psychological corollary, and since the distinctive features of a response have not been, perhaps cannot be, settled with anything like the firmness that linguistis have achieved for the distinctive features of the phoneme. In spite of these discrepancies, however, one can think of several famous psychologists who might have written our hypothetical critique.

The distinctive features of the phoneme are, according to Jakobson and Halle, binary in nature: vocalic/non-vocalic, consonantal/non-consonantal, compact/diffuse, tense/lax, voiced/voiceless, nasal/oral, discontinuous/continuous, strident/mellow, checked/unchecked, grave/acute, flat/plain, and sharp/plain. These twelve binary attributes are adequate to describe all the phonemes in any language. The phonemic code is acquired by children in a regular order as they successively master each of the distinctive features employed in their own language. The same order is reversed in the deterioration of speech observed in aphasia.

The second half of the book consists of an essay by Jakobson on two aspects of language and two types of aphasic disturbance. The two aspects are *selection* and

combination; a linguistic unit must be selected from a set of alternatives and must be combined with others to form a more complex unit. One type of aphasic disorder relates to selection, the other to combination. This is not a new observation, but what is new is the emphasis that Jakobson puts on the dichotomy and the extent to which he develops its implications. Selection depends upon the ability to conceive the relation of similarity, and when it is damaged the metaphorical aspects of language suffer. Combination depends upon the ability to conceive the relation of contiguity, and when it is damaged the metonymic aspects of language suffer. Similarity and contiguity are certainly familiar terms to a psychologist, dignified as they are by a long philosophical history as laws of association and by their importance for modern theories of learning. The relation of Jakobson's selection-similarity-metaphor and combination-contiguity-metonym syndromes to current psychological theorizing about similarity and contiguity is worth looking into, whether or not this classification gains favor with the clinicians who deal with aphasic patients.

Psychologists who are willing to take their ideas where they find them will enjoy this little book. It requires some recoding to get it into our psychological jargon, but the labor is not excessive in view of the bounty bestowed.

Harvard University   GEORGE A. MILLER

Group Processes: Transactions of the First Conference. Edited by BERTRAM SCHAFFNER. New York, Josiah Macy Jr. Foundation, 1955. Pp. 334.

Frank Beach assumes "that to understand group processes we need considerable understanding of the individuals who make up the group." This view apparently served as the rationale for the conference of psychologists and others professionally concerned with animal behavior which is reported in this volume. A good part of it is devoted to exchanges among representatives of European ethology and American psychology, exchanges which often serve to emphasize the differences of approach and program between the two disciplines. Of lesser volume—probably because this is the first conference of a scheduled series—is the contribution of those trying to relate animal behavior to the more clinical and social side of psychology. The Table of Contents listing chapter titles and initiating discussants indicates the focus of interest: *Ontogeny and Living Systems,* by Frank Beach; *Psychology and Ethology as Supplementary Parts of a Science of Behavior,* by Niko Tinbergen; *Morphology and Behavior Patterns in Closely Allied Species,* by Conrad Lorenz; *Dynamics of the Mother-Newborn Relationship in Goats,* by Helen Blauvelt; *The Perception of Animal Behavior,* by Daniel Lehrman; *Group Processes in Lower Vertebrates,* by L. Thomas Evans.

Despite the shortcomings of this form of publication, so pungently aired by Miller in a review of a similarly printed conference (this JOURNAL, 66, 1953, 661-663), those interested enough to read (sometimes between the lines) in this one may gain an unusually vivid perspective of one kind of psychological thinking entrenched among Americans. One may wonder indeed whether Miller's jocose sarcasm blended, curiously enough, with a stern plea for responsible scientific communication, does not overlook certain values of face-to-face discussion available even in these less-than-optimal presentations. Fundamental differences in assumption and motive appear much more sharply in such informal discussion, with its ques-

tions and answers, than in the stylized and edited writing of the standard sources. Shop-talk, which may be so familiar as to bore the experts of a field, is simply not available to the majority of the potentially interested. Besides, may not the printed conference be a modern substitute for the vanishing art of discursive correspondence and hence a similar kind of source material for future historians? The following is a sample of this aspect of the book.

Tinbergen deals directly with the question of ethology (pronounced with a short *e*, we learn) versus psychology. He speaks first of the similarity of approach of the two disciplines, which amounts to the use of scientific method. More interesting are the differences that he enumerates and the accompanying discussions. Most provocative, as demonstrated by the participants' responses, is his assertion that the ethologists emphasize the problem of survival-value and evolution, which the psychologist tends to ignore in his more exclusive concern with "immediate causation of the observed phenomena." Following Tinbergen's presentation, Lorenz apparently feels obliged to defend the causal nature of processes resulting in "survival value," and he ends a long statement by reassuring his audience that, "I know that we really are quite as 'causalistic' as any of the American school." If Lorenz' motive for making this confession and plea was not already clear, it was soon clarified in the course of an exchange with Schneirla following the latter's veiled accusation of the evils ("teleology" and "purpose") inherent in ethological thinking. Lorenz' response is a strong endorsement of Schneirla criticism of the teleology that implies a "directing factor" as distinct from the ethologist's view of the direction of life-processes, but Schneirla, apparently not convinced, continues his argument for objectivity in the study of behavior. At this point Spiegel, a psychiatrist, remarks, "I have been sitting on the periphery of this discussion and, really, a large amount of the emotional force of the difference in view is passing me by." We may assume that he has been struck by the vehemence of these oblique criticisms and rejoinders. Previously, conference-guest Mayr had astutely dichotomized biologists (he is one) into functionalists—those who ask "how" a process operates here and now—and evolutionists—those who ask "why" in the phylogenetic sense—thereby supplying a neat and neutral way of understanding the difference. To judge from subsequent discussion, alas, Mayr's sensible view was disregarded, and it becomes clear to the reader, as well as to some of the participants, that this is not merely a casual misunderstanding.

Other basic oppositions can be sampled from the discussions. The ethologists show themselves committed to "innate motor coordinations," to "innate releasing mechanisms," and they point out inadequacies in reflex and conditioned-reflex interpretations of behavior, but only in the face of constant criticism from the psychologists who object to postulated *CNS*-mechanisms, demand that interpretations of behavior be made first in terms of peripheral mechanisms, and continually raise the question of the role of experience in development. Once again, the adumbration of these issues beyond all logical need points to basic differences of motive and heritage. Mayr's explanation of the strongly functional and experimental bias of American psychologists by the fact that many work closely with physiologists and medical men seems inadequate. After all, we know that American psychology has been enormously influenced by ideas of biological evolution and hence we are left to explain why at present these two products of evolutionary

thinking have themselves evolved with different *fragestellungen.* Certain reasons
suggest themselves and their cogency can be roughly gauged by the issues raised
in this conference.

The American psychologists of the turn of the century had already a "tempered
Darwinism," in the words of Boring. Turning from the doctrine of early compara-
tive psychologists regarding mind as another organ evolved by environmental se-
lection (propaganda for the theory of evolution), they became concerned about
the ways in which the individual acts upon and transforms his environment, with
the upwardly-mobile American male as the touchstone for psychological theory.
Ontogenetic adaptation—learning by means of which the individual comes to know
his world—was a central feature of the process. The early and radical Thorndike
made a significantly-qualified bow to evolution by equating the innate component
of intelligence with the number of potential *associations* to be formed in the
development of the individual. To this basic concern with the transactions between
organism and environment, the concept of conditioned reflex was one welcome
device among others for causal explanation in terms only of events at the peripheral
junction of the two, and without the vagaries of mental states or internal mecha-
nisms.

It is not surprising that to the inheritors of this complex of ideas, the morphologi-
cal approach of the ethologists—focusing of necessity on rigid species-specific
characters (usually in lower animals) for which the methodology of discovery de-
mands an environment theoretically held constant—is threatening in its nearly com-
plete disregard of the major programmatic concerns of the American psychologists.
To this may perhaps be added the distaste resulting from a tendency in America to
identify hereditarian with anti-democratic views. Although these substantial di-
vergences of approach exist. they do not necessarily preclude a useful outcome of
the new interaction exemplified by this conference. Both sides may as a consequence
be forced to clarify the nature and bounds of their concepts and thereby improve
a state of affairs under which Tinbergen can say, "Innate behavior is behavior that
has not been changed by learning processes," while Beach echoes him with slightly
different emphasis: "Instinct is what is not learned." Perhaps, as Beach seems to
believe, the presently-popular study of the role of early experience in development
may provide a common ground for working out relations of the precision required
to define and distinguish independent subject-matters.

Brandeis University   RICHARD HELD

Basic Statistical Concepts. By JOE KENNEDY ADAMS. New York, McGraw-Hill
Book Company, 1955. Pp. xvi, 304.

The Essentials of Educational Statistics. By FRANCIS G. CORNELL. New York,
John Wiley & Sons, 1956. Pp. xii, 375.

The contrast between these two books provides a sharp commentary on changes
which have occurred during the last twenty years in the writing of statistics texts
by and for people in psychology and allied fields. Cornell has aimed his book ex-
clusively at advanced students of education, while Adams will, if he can, take
the entire undergraduate student body as his audience. Cornell doubts that his
prospective readers are adequately prepared in arithmetic and algebra to grapple
with the contents of his book, and he urges them not to despair of acquiring the

necessary background. Adams compliments his audience not only by his confidence that they already can do arithmetic and know some algebra, but also by making bold to assert that they can be taught sufficient calculus to deal adequately with the elementary concepts of mathematical statistics.

In a sense, Adams has really produced two books between the same covers. One is the 211 pages of text, in which he discusses with great clarity, succinctness, and elegance the basic ideas of statistics (including a brief chapter on non-parametric methods). He discusses them as mathematical propositions. He makes such good use of the power of general conceptions, that a student could go from his book to one of the references in mathematical statistics he cites (Cramer, Hoel, and Wilks) without feeling he has been transported to another universe of discourse. The second Adams 'book' is his Appendix B (pp. 219-254), in which he presents discussions and proofs of many of the propositions treated in the text. The complete list of topics is too long to cite here, but a few examples will convey the richness of his offerings: limits of sequences and series, integration and differentiation, mathematical expectation, moment-generating functions, Tchebysheff's inequality. The text and Appendix B complement each other. Only once in the text (p. 75) does Adams state a theorem without either (*a*) proving it on the spot, or (*b*) telling the reader that a proof is in Appendix B, or (*c*) referring him to a proof given by another author. On the other hand, Appendix B could not stand in its present form without the development offered in the text.

Examples and exercises are presented in large numbers, but they are simple, direct, and sharply limited to the points to which they are relevant. Another appendix provides correct answers for the odd-numbered exercises.

In Adams' book the reader will look in vain for reference to the history of statistics, or its *special* applications. Time-series, index numbers, psychological-test statistics, and so on, are not discussed, although the reader is left in no doubt about the range of statistical applications, for the exercises *nominally* cover every relevant field from Agronomy to Zoology.

Adams' whole conception of how to introduce statistics is the significant feature of the changing aspect of statistics texts which he represents. Statistics is a field of broad applicability and it possesses an intrinsically interesting and rational structure of *general* ideas. The book lays bare these ideas and a good bit of the structure which connects them. Adams disclaims that his book is unique in these respects, but it is certainly a member of a very small class.

Cornell's book also shows the symptoms of change, but the basic pattern of the book is little different from the texts of twenty years ago. There is some new material inserted among the old, but the change is that of accretion to an established structure. It is a sound and workmanlike job, and it is best when Cornell forgets the highly specialized nature of his audience and concentrates on the statistics. The writing sometimes evokes the image of a Scoutmaster leading a hike and frequently looking over his shoulder to see how many of his charges are still following him. The presentation, he says, is "nonmathematical" but he dismisses the notion that statistics can genuinely be understood without mathematics. In fact, among those who use the same mode of presentation Cornell get closest to revealing some of the logic behind the assertions about statistical procedures and terminology. He obviously does not intend that his book will find wide use beyond

students of education, and that intention will very probably be realized. His examples and exercises are drawn almost exclusively from educational data, and hence will not compete in interest with the examples in equally parochial books long used in psychology. His discussions of distribution-theory are one JND better than those of any comparable book, but the above considerations still determine against its use in other fields.

On the physical side, both books are attractive. Adams' publisher has provided the larger type, but this reviewer had difficulty in resolving some of Adams' tiny subscripts and superscripts in the usual reading posture. Authors of books on mathematical subjects should hold a general conference with all book-designers to attempt to reduce the inverse correlation between the complexity of equations and the size of type used in them. Otherwise the day will come when all their readers will be holding individual conferences with eye-specialists. Typographical errors appear to be rare in these books. This reviewer found none in Adams' and only two in Cornell's, although both might give trouble: the bar indicating a general mean has been left off one of the Xs in equation 12.16, and also in the first paragraph on page 280.

Each of the authors has accomplished his stated objectives. Because of the restrictions Cornell has placed on his audience, people interested in training research workers outside the field of education will find his book of little help. Because Adams represents a new departure, it will be the rare scientist, doubling as a teacher of statistics, who will be able to help students get all that is to be gotten from his book, but happy will be those trainers of scienists whose students have mastered such a book before they begin graduate study.

University of Pennsylvania   JAMES C. DIGGORY

The Power Elite. By C. WRIGHT MILLS. New York, Oxford University Press, 1956. Pp. 423.

The earnest thinking and reappraisal of the problems of political democracy following on the heels of the first world war gave us Walter Lippmann's *Public Opinion.* *The Power Elite* is the volume appearing in the period after the second world war most likely to be remembered for providing the sound social and political insights so necessary if we are to keep our political perceptions abreast of the rapid pace with which events are transforming the social facts we belatedly perceive. Since only in the now-fictitious, fairy-tale, 18th-century democracy does the public know facts and arrive at decisions, Professor Mills sets himself the task of discovering precisely who today make the decisions that have major consequences for the political, economic, and military future and fate of the United States. The decision-makers he calls the "power elite," for they are to be found at the apex of the power-hierarchy in each of the specified areas of national life.

The appearance of this volume, just when we are beginning to learn something of the fantastic consequences of individual decisions at the apex of the power-pyramid in Russia, could hardly have been more appropriately timed. Professor Mills does not champion a theory of history as revolving about specific individuals, yet he recognizes (and it is high time we all did) that we are moving into a period of history in which the potentialities inherent in decisons made at the top

of the power-ladder are increasingly significant for every man-jack of us. The myth of the impersonal character of historical events is becoming dangerous. The wielders of power much prefer to hide their decisions and machinations behind the anonymity of Manifest Destiny, the Will of the People, or the National Will, but no social science worthy of the name, not even a respectable history, can longer take refuge in such dodges. Professor Mills asks point-blank where critical decisions are made and who makes them. He finds the answers in none of the traditional concepts of Class, nor of the Sixty Families, nor in a Managerial Revolution. Instead, he points to the big rich and to those in control of the giant corporations, to the higher echelons of the political hierarchy, and the top levels of the military bureaucracy (all three of which he finds increasingly interlocking and interchangeable in personnel).

Who are to be found in the positions at the apex of the power-pyramids in each of these three dominant social dimensions? Here we have clearly pointed out to us their origins in the 'correct' private schools and colleges, their far-above-chance derivation from proper families, businesses, professions, political, and religious allegiances. Yet the gates are not wholly closed to those who successfully overcome the handicap of not having the proper origins, and who, by the very force of traits and accomplishments, come to hew out a place of power for themselves. No degree of 'correct' origin can wholly substitute for the qualities necessary for the extension of a power-area in a world in which power-relations are always fluid, but origin does have an impressive total effect on the final statistics of attainment to the power positions.

The author emphasizes, and correctly, the changing status of the various power-hierarchies in the current social scene. The depression-thirties put a quite different pattern of personalities into the top echelons of the political power-structure which was then clearly playing the dominant role, what with Big Business being in the doghouse because of the business-slump and unemployment. The recent wars have, for the first time in a period of peace, put the military brass into the top power-trinity, where they are putting their imprint most profoundly upon the whole American culture. They have sold us the military definition of reality; they are riding high, and (not to be forgotten) they step readily into the top levels of both the political and industrial world in a culture increasingly dominated by military perceptions, values, and arms-contracts. The economic prosperity resulting from federal spending has given us a crop of political office-holders whose values and points of view adjust easily to those of the military mind.

No review could possibly do justice to the author's masterly analyses of the various elites and their interlocking relationships: The Metropolitan 400, The Corporate Rich, The Very Rich, The Warlords, and the Political Directorate. The Celebrities are depicted as providing safe and innocuous distractions for the Mass Society. Their chummy relationships with the various elites, their role in catalyzing a mass-acceptance of the values of the elite, are clearly and bitingly set forth. The old 'theory of balance' as an account of the ruling process in a democratic society is dismissed with the verdict that it is useful today only as a moral device invoked by privileged groups to justify and to maintain their dominant positions. In their retreats in the face of assaults by political primitives, educators and liberals have

been reduced to defending liberties in the abstract instead of practicing them in the concrete. There is more than a little basis for that charge.

The chapters on The Conservative Mood and The Higher Immorality are as effective polemical writing as has appeared in many a day. The thesis advanced is that a demagogic and stupid conservatism is riding high and undoing as many of the liberal gains of the pre-war decade as possible. This conservatism glorifies its own materialism while denouncing that of the communistic enemy. Such organization as it has is directed to the promotion of its own vested interests, giving us not an intelligent social leadership but an "organized irresponsibility." The volume should be required reading for all college administrative officers.

New York University   GEORGE B. VETTER

Mysterium Coniunctionis: Untersuchung über die Trennung und Zusammensetzung der seelischen Gegensätze in der Alchemie. I. By C. G. JUNG with the collaboration of M.-L. VON FRANZ. Zurich, Rascher Verlag, 1955. Pp. xv, 284.

Studien zur analytischen Psychologie C. G. Jungs. From the C. G. JUNG INSTITUTE. Zurich, Rascher Verlog, 1955. Vol. 1: *Beiträge aus Theorie und Praxis;* pp. xi, 396. Vol. 2: *Beiträge zur Kulturgeschichte;* Pp. ix, 397.

Anyone interested in a panoramic view of Jungian activities in the past decade would be very likely to find it in these three tomes. This is not meant to imply that the task would be a simple one, or one that would not demand some previous preparation on the part of the reader. Within these pages, however, one would be exposed to the scholarly efforts of Jung himself and to most of the leading proponents of Analytical Psychology.

The major concepts of his theoretical position were outlined by Jung many years ago. While they grew primarily out of his observations as a psychotherapist, his efforts in the later years of his life have been concerned with the documentation of these views by examining such diverse subject matters as mythology, religion, and alchemy. He finds mirrored in these several realms the archetypal process of individuation, the ultimate goal of psychic life.

Mysterium Coniunctionis is a work in three parts dedicated to the task of examining alchemy and its problems, not as a prelude to a scientific chemistry, but as a philosophy. The purpose, however, is not to add to the vast collection of esoteric works that appear on library shelves, but rather to relate the findings of such an investigation to the understanding of the more distant and universal aspects of man's unconscious. Part I, which is all that has been published thus far, deals with the specific problem, central to Jungian thought, of "opposites," their representation, personification, paradoxical nature, separation, and unification. Excursions are made into the symbolism of alchemy for the various forms that the opposites can take, for the substances from which they can emerge, and for the conditions under which they are differentiated and reunited. Parallels are drawn between the representations found in the alchemical literature and those found in the mystical writings of the Gnosis, Cabbala, and the ritual of formal religion. This is done to substantiate the view that these little-regarded, abandoned literatures are nothing more than projections of man's collective unconscious and as such are veritable storehouses of knowledge about man's psychic life.

How does one evaluate such a book? Its subject matter and method of inquiry are foreign to the psychological milieu of this country. It asks for giant steps in the acceptance of its assumptions, steps that can appear ludicrous upon encounter. Shall one use the criterion of logic, or originality, or credibility? This reviewer can find no simple solution and perhaps his difficulty is but a specific example of the more general dilemma of evaluating Jung's contributions to psychology. This is a scholarly document, imaginative and resourceful in its approach. It is not a book for the uninitiated. One should be steeped in both Jungian language and ideas before making its acquaintance.

The two volumes of the *Studien* contain individual papers, in English, German, and French, written by almost all of the people who have appeared as lecturers at the C. G. Jung Institute of Zurich since it was founded in 1948. The work was offered as a tribute to Jung on the occasion of his eightieth birthday, celebrated in the summer of 1955. The essays are grouped under three headings: those related to theory, those growing primarily out of case-material, and those concerned with cultural history. Some papers are elaborations of Jung's basic ideas, some attempt to confirm their validity. A few deal with unsolved problems in the system, while others establish relationships with subject matters not heretofore considered, *e.g.,* projective testing. The range of topics discussed in extremely broad. Within these pages one can find excursions into the writing of Shakespeare, the mythology and dreams of American Indians, the classical dance, fairy tales, various religions, psychosomatics, the history of medical psychology, and so forth.

There is much that is interesting, informative, and provocative in these books. They stand as a fine tribute to Jung's pioneering efforts. Their presentation of modern Jungian thought can be of value to American psychologists.

Princeton University IRVING E. ALEXANDER

Pharmacologic Products Recently Introduced in the Treatment of Psychiatric Disorders. Edited by WILLIAM T. LHAMON. Washington, American Psychiatric Association, 1955. Pp. 152.

From time immemorial men have sought magic elixirs and potions to ease troubled minds, and it is surprising indeed that their haphazard explorations, attended by necromacy and primitive folk-medicine, should have been so successful. (Most of the agents causing psychological changes are of vegetable origin and belong to the class of substances known as *alkaloids,* with the outstanding exception of that universal sorrow-solvent, *ethyl alcohol.*) Recent years have witnessed a sudden surge of interest in the behavioral effects of drugs in general and of certain "psychotomimetic" and "tranquilizing" drugs in particular. Hence it is not surprising that the American Psychiatric Association should have devoted the first of a new series of research reports to a conference on psychotherapeutic drugs.

The conference was divided into three symposia. The first dealt with somewhat more basic scientific questions—physiological, pharmacological, neurological, and psychological—of drug action. The second dealt with the "clinical value and limitations of chlorpromazine," and the third with the "clinical value and limitations of reserpine, Meratran, and 'a new blocking agent.' " It is likely that the experimental psychologist would find more to interest him in the first symposium

than in the more clinically oriented second and third. In general, a reader of this JOURNAL would find these reports somewhat disappointing with respect to method, which is not surprising when one considers that of the 29 contributors only two, N. S. Kline and O. D. Murphree are listed in the 1955 directory of the American Psychological Association. (The majority are psychiatrists.) The authors of the only study giving significance-levels (Gaitz and others) report that the Lorr Scale (for rating psychotics) showed no significant difference between patients receiving chlorpromazine and those receiving "standard" hospital therapies, yet they conclude that chlorpromazine is an effective agent for hospitalized psychotic patients.

The first paper is a preliminary report by Schneider on some effects of amobarbital, reserpine, and chlorpromazine on the Funkenstein test. The findings here are certainly suggestive of an effect of these drugs upon the blood-pressure regulating mechanisms, but the small number of Ss and the variability of the results (of six Ss given chlorpromazine, three showed increased "central sympathetic reactivity" and three showed decreased CSR) make the studies rather inconclusive, except to point out a more consistent effect of amobarbital than of chlorpromazine and reserpine upon CSR. Whether the author's hypothesis that reserpine normalizes the CSR is correct remains to be proven. The second paper, by Cohen and Nash, presents some interesting data (only of value to psychologists who go out into the sunshine) that chlorpromazine may photosensitize certain rare individuals. This is really a clinical report, and the authors frankly admit that "no definitive conclusions can be drawn." Ayd's paper on the physiological and neurological action of chlorpromazine is a scholarly and well-documented review of the subject. It is easy to agree that "chlorpromazine is a truly remarkable drug, which besides its therapeutic properties, should contribute much to our knowledge and understanding of the physiologic and neurologic foundations of human behavior." This hope will be realized, of course, only if work in such areas is coördinated with controlled, quantitative behavioral studies.

Monroe, Heath, Mickle, and Miller report on cortical and subcortical recordings in humans and monkeys with the administration of reserpine, chlorpromazine, and pentylenetetrazol (Metrazol). Their findings of an alert type of electrical record in a somnolent-appearing animal have since been confirmed by others. It is interesting that hippocampal and septal leads appear to show abnormal electrical activity both in naturally occurring psychotic states and in those induced by LSD-25. Further work conducted with due humane consideration will be needed to confirm this important finding. West sets forth nine excellent rules to guide research on the behavioral effects of drugs, and it is regrettable that they were not applied more often in the research reported in this volume. Basically, good research in this field must follow the same rules as that in any field: adequate controls, statistical evaluation of results, and legitimate conclusions. For psychiatrists to work profitably within this framework it is clear that reliable objective methods of evaluating psychopathology must be devised. This is a problem much more central and pressing than the search for new drugs to treat mental illness, although one of the most promising contributions of the burgeoning field of psychopharmacology (hinted at by this collection) is that it will undoubtedly stimulate psychiatrists and their colleagues to create such methods.

Albert Einstein College of Medicine   M. E. JARVIK

INDEX

By H. W. Stevenson, University of Texas

Authors

(The names of authors of original articles are printed in Capitals; of authors of books reviewed, in roman; and of reviewers, in *italics*.)

(References in *italic* figures are to reviews.)